

UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
CURSO DE GRADUAÇÃO EM ENGENHARIA DE TELECOMUNICAÇÕES

Jerônimo Acker D'Ornellas

ESTUDO DA MODELAGEM ESTATÍSTICA EM IMAGENS SAR

Santa Maria, RS
2023

Jerônimo Acker D'Ornellas

ESTUDO DA MODELAGEM ESTATÍSTICA EM IMAGENS SAR

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Telecomunicações, Área de Concentração em Engenharia IV, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Bacharel em Engenharia de Telecomunicações**.

ORIENTADOR: Prof. Natanael Rodrigues Gomes

Santa Maria, RS
2023

©2023

Todos os direitos autorais reservados a Jerônimo Acker D'Ornellas. A reprodução de partes ou do todo deste trabalho só poderá ser feita mediante a citação da fonte.

End. Eletr.: jeronimo_acker@hotmail.com

Jerônimo Acker D'Ornellas

ESTUDO DA MODELAGEM ESTATÍSTICA EM IMAGENS SAR

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Engenharia de Telecomunicações, Área de Concentração em Engenharia IV, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Bacharel em Engenharia de Telecomunicações**.

Aprovado em 10 de julho de 2023:

Natanael Rodrigues Gomes, Dr. (UFSM)
(Presidente/Orientador)

Candice Muller, Dra. (UFSM)

Samuel Tumelero Valduga, Dr. (UFSM)

Santa Maria, RS
2023

RESUMO

ESTUDO DA MODELAGEM ESTATÍSTICA EM IMAGENS SAR

AUTOR: Jerônimo Acker D'Ornellas

ORIENTADOR: Natanael Rodrigues Gomes

As imagens SAR se tornaram ferramentas importantes com aplicações em áreas como monitoramento ambiental e segurança nacional. A modelagem estatística é uma etapa importante no processamento e análise de imagens SAR e seus resultados são amplamente usados em áreas de pesquisa como detecção de alvos e classificação de imagens. Neste estudo são exploradas cinco funções de distribuição de probabilidade: K, G0, Rayleigh, Lognormal e Weibull para modelar regiões de veículos na floresta, lago, floresta e estrada. Os dados coletados dessas regiões pela aeronave CARABAS-II em frequência VHF (20-90 MHz) são analisados quanto ao número de espalhadores e distribuições que melhor descrevem as regiões estudadas. Os parâmetros de cada distribuição são estimados com base no Método da Máxima Verossimilhança e o Método dos Momentos e as distribuições são comparadas utilizando a distância KS como métrica.

Palavras-chave: Modelagem estatística, Imagem SAR, CARABAS-II, Número de espalhadores.

ABSTRACT

STUDY OF STATISTICAL MODELING IN SAR IMAGES

AUTHOR: Jerônimo Acker D'Ornellas

ADVISOR: Natanael Rodrigues Gomes

SAR images have become important tools with applications in areas such as environmental monitoring and national security. Statistical modeling is an important step in SAR image processing and analysis, and its results are widely used in research areas such as target detection and image classification. In this study, five probability distribution functions are explored: K, G0, Rayleigh, Lognormal, and Weibull to model vehicle regions in the forest, lake, forest, and road. The data collected from these regions by the CARABAS-II aircraft at VHF frequency (20-90 MHz) are analyzed for the number of scatterers and distributions that best describe the studied regions. The parameters of each distribution are estimated based on the Maximum Likelihood Method and the Method of Moments, and the distributions are compared using the KS distance as a metric.

Keywords: Statistical Modeling, SAR image, CARABAS-II, Number of scatterers.

LISTA DE FIGURAS

Figura 2.1 – Diagrama de blocos dos componentes básicos de um sistema de radar.	13
Figura 2.2 – Geometria de imageamento SAR. r_o é a menor distância de aproximação, θ_a é a largura do feixe do azimute e v é a velocidade do sensor fixado na plataforma.	15
Figura 2.3 – Modo de aquisição de imagens Stripmap.	17
Figura 2.4 – Modo de aquisição de imagens Spotlight.	18
Figura 2.5 – Modo de aquisição de imagens Scan.	19
Figura 2.6 – Representação do pixel no formato complexo.	20
Figura 2.7 – Variações de retroespalhamento.	21
Figura 2.8 – Imagem SAR com resolução 1m x 1m na banda de frequência Ku com as variações pixel a pixel no tom da imagem características do speckle.	22
Figura 2.9 – Exemplo distribuição simétrica.	24
Figura 2.10 – Exemplo distribuição com assimetria negativa.	25
Figura 2.11 – Exemplo distribuição com assimetria positiva.	25
Figura 3.1 – Diagrama de blocos do processo de modelagem estatística paramétrica.	28
Figura 4.1 – Imagem aérea da região de teste dentro da localidade RFN Vidsel.	38
Figura 4.2 – Regiões com diferentes espalhamentos selecionadas a partir do conjunto de dados CARABAS-II para análise do processo de espalhamento e das distribuições com melhor ajuste. (a) Veículo Terrestre, (b) Estrada, (c) Floresta, (d) Lago.	39
Figura 4.3 – Diagrama de blocos das etapas desenvolvidas no trabalho.	40
Figura 5.1 – Exemplo de imagem analisada da região dos veículos na floresta.	41
Figura 5.2 – Histograma e PDF das distribuições da região dos veículos da imagem do <i>deployment</i> Sigismund.	45
Figura 5.3 – CDF da região dos veículos na floresta da imagem do <i>deployment sigismund</i>	45
Figura 5.4 – Exemplo de imagem do <i>deployment</i> Fredrik para a análise da região da floresta.	46
Figura 5.5 – Comparação das distribuições da região da floresta	50
Figura 5.6 – CDF da região da floresta da imagem do <i>deployment sigismund</i>	51
Figura 5.7 – Imagem do <i>deployment</i> Adolf-Fredrik para análise da região do lago.	52
Figura 5.8 – Comparação das distribuições da região do lago	55
Figura 5.9 – CDF da região do lago da primeira imagem do <i>deployment Adolf-Fredrik</i>	56
Figura 5.10 – Imagem do <i>deployment</i> Karl para análise da região da estrada.	57
Figura 5.11 – Comparação das distribuições da região da estrada.	60
Figura 5.12 – CDF da região da estrada da primeira imagem do <i>deployment Fredrik</i>	61
Figura 6.1 – Comparação do número de espalhadores N do estudo de referência e N_m sendo o número de espalhadores médio estimados no estudo.	62

LISTA DE TABELAS

Tabela 2.1 – Bandas de frequência, faixas de comprimento de onda e frequências associadas comumente usadas em sistemas SAR.....	14
Tabela 3.1 – PDFs das distribuições e métodos de estimação de parâmetros usados no trabalho.....	34
Tabela 4.1 – Parâmetros usados no sistema de radar CARABAS-II,.....	37
Tabela 5.1 – Número de espalhadores da região dos veículos do deployment Sigismund.....	42
Tabela 5.2 – Distância KS das distribuições para a região dos veículos na floresta ...	42
Tabela 5.3 – Número de espalhadores da região dos veículos do deployment Karl ...	42
Tabela 5.4 – Distância KS para a região dos veículos do deployment Karl.....	43
Tabela 5.5 – Número de espalhadores da região dos veículos Fredrik.....	43
Tabela 5.6 – Distância KS para a região dos veículos do deployment Fredrik.....	43
Tabela 5.7 – Número de espalhadores da região dos veículos do deployment Adolf-Fredrik.....	44
Tabela 5.8 – Distância KS para a região dos veículos do deployment Adolf-Fredrik ...	44
Tabela 5.9 – Número de espalhadores da região da floresta Sigismund.....	47
Tabela 5.10 – Ajuste de distribuições usando a distância KS do estudo de referência..	47
Tabela 5.11 – Número de espalhadores da região da floresta Karl.....	47
Tabela 5.12 – Distância KS para a região da floresta do deployment Karl.....	48
Tabela 5.13 – Número de espalhadores da região da floresta Fredrik.....	48
Tabela 5.14 – Distância KS para a região da floresta do deployment Fredrik.....	48
Tabela 5.15 – Número de espalhadores da região da floresta.....	49
Tabela 5.16 – Distância KS para a região da floresta do deployment Adolf-Fredrik. ...	49
Tabela 5.17 – Número de espalhadores da região do lago <i>deployment</i> Sigismund....	52
Tabela 5.18 – Distância KS das distribuições para a região do lago do <i>deployment</i> Sigismund.....	53
Tabela 5.19 – Número de espalhadores da região do lago do <i>deployment</i> Karl.....	53
Tabela 5.20 – Distância KS das distribuições para a região do lago do <i>deployment</i> Karl.....	53
Tabela 5.21 – Número de espalhadores da região do lago do <i>deployment</i> Fredrik....	54
Tabela 5.22 – Distância KS das distribuições para a região do lago do <i>deployment</i> Fredrik.....	54
Tabela 5.23 – Número de espalhadores da região do lago do <i>deployment</i> Adolf-Fredrik	54
Tabela 5.24 – Distância KS das distribuições para a região do lago do <i>deployment</i> Adolf-Fredrik.....	55
Tabela 5.25 – Número de espalhadores da região da estrada para o <i>deployment</i> Sigismund.....	57
Tabela 5.26 – Distância KS das distribuições para a região da estrada do <i>deployment</i> Sigismund.....	58
Tabela 5.27 – Número de espalhadores da região da estrada para o <i>deployment</i> Karl.	58
Tabela 5.28 – Distância KS das distribuições para a região da estrada do <i>deployment</i> Karl.....	58
Tabela 5.29 – Número de espalhadores da região da estrada para o <i>deployment</i> Fredrik.....	59
Tabela 5.30 – Distância KS das distribuições para a região da estrada do <i>deployment</i>	

Fredrik.	59
Tabela 5.31 – Número de espalhadores da região do lago do <i>deployment</i> Adolf-Fredrik.	59
Tabela 5.32 – Distância KS das distribuições para a região da estrada do deployment Adolf-Fredrik.....	60
Tabela 6.1 – Média da distância KS das distribuições do trabalho de conclusão de curso.....	63
Tabela 6.2 – Ajuste de distribuições usando a distância KS do estudo de referência. .	63

LISTA DE ABREVIATURAS E SIGLAS

<i>PRF</i>	Frequência de repetição de impulsos
<i>RCS</i>	Seção transversal do radar
<i>MLE</i>	Estimativa por máxima verossimilhança
<i>MoM</i>	Método dos momentos
<i>PDF</i>	Função densidade de probabilidade
<i>CDF</i>	Função distribuição acumulada

SUMÁRIO

1	INTRODUÇÃO	11
1.1	OBJETIVO	11
1.2	ORGANIZAÇÃO DO TRABALHO	12
2	FUNDAMENTAÇÃO TEÓRICA	13
2.1	RADAR DE ABERTURA SINTÉTICA	13
2.2	IMAGEAMENTO SAR	14
2.3	TÉCNICAS DE IMAGEAMENTO SAR	17
2.4	TIPOS DE IMAGENS SAR	19
2.5	SINAL RETROESPALHADO	20
2.6	SPECKLE	21
2.7	ESTATÍSTICA DESCRITIVA	22
2.7.1	Medidas de tendência central	22
2.7.2	Medidas de dispersão	23
2.7.3	Formato dos dados	24
3	MODELAGEM ESTATÍSTICA DE IMAGENS SAR	27
3.1	INTRODUÇÃO	27
3.2	MODELAGEM ESTATÍSTICA BASEADA NO MODELO DO ESPALHADOR COERENTE	28
3.2.1	Modelo Rayleigh do speckle	29
3.2.2	Modelagem estatística baseada no modelo speckle NR	30
3.2.2.1	<i>Distribuição K</i>	31
3.2.2.2	<i>Distribuição G^0</i>	31
3.3	MODELAGEM ESTATÍSTICA BASEADA NOS MODELOS EMPÍRICOS	32
3.3.1	Distribuição Lognormal	32
3.3.2	Distribuição Weibull	32
3.4	ESTIMAÇÃO DOS PARÂMETROS	32
3.4.1	Método dos momentos	33
3.4.1.1	<i>Estimação dos parâmetros da distribuição K</i>	33
3.4.2	Método da máxima verossimilhança	33
3.4.3	Estimação dos parâmetros para cada distribuição	34
3.5	RELAÇÃO DO NÚMERO DE ESPALHADORES E O MODELO DO SPECKLE	35
3.6	MÉTRICA DE AVALIAÇÃO	35
3.6.1	Teste de Kolmogorov-Smirnov	35
4	MATERIAIS E MÉTODOS	37
4.1	DESCRIÇÃO DA BASE DE DADOS	37
4.1.1	CARABAS-II	37
4.1.2	Conjunto de dados	38
4.1.3	Metodologia	39
5	RESULTADOS	41
5.1	MODELAGEM ESTATÍSTICA DA REGIÃO DOS VEÍCULOS NA FLORESTA ..	41
5.1.1	Deployment Sigismund	42
5.1.2	Deployment Karl	42
5.1.3	Deployment Fredrik	43
5.1.4	Deployment Adolf-Fredrik	44
5.1.5	Visualização das distribuições da região dos veículos na floresta	44

5.2	MODELAGEM ESTATÍSTICA DA REGIÃO DA FLORESTA	46
5.2.1	Deployment Sigismund	47
5.2.2	Deployment Karl	47
5.2.3	Deployment Fredrik	48
5.2.4	Deployment Adolf-Fredrik	49
5.2.5	Visualização das distribuições da região da floresta	49
5.3	MODELAGEM ESTATÍSTICA DA REGIÃO DO LAGO	51
5.3.1	Deployment Sigismund	52
5.3.2	Deployment Karl	53
5.3.3	Deployment Fredrik	54
5.3.4	Deployment Adolf-Fredrik	54
5.3.5	Visualização das distribuições da região do lago	55
5.4	MODELAGEM ESTATÍSTICA DA REGIÃO DA ESTRADA	56
5.4.1	Deployment Sigismund	57
5.4.2	Deployment Karl	58
5.4.3	Deployment Fredrik	59
5.4.4	Deployment Adolf-Fredrik	59
5.4.5	Visualização das distribuições da região da estrada	60
6	ANÁLISE DOS RESULTADOS	62
6.1	NÚMERO DE ESPALHADORES	62
6.2	MODELAGEM DAS DISTRIBUIÇÕES	63
7	CONCLUSÃO	65
	REFERÊNCIAS BIBLIOGRÁFICAS	66

1 INTRODUÇÃO

O rápido avanço de tecnologias de radar de abertura sintética (SAR, do inglês *Synthetic Aperture Radar*) e a ampla aquisição de imagens SAR fez com que a interpretação das imagens registradas por estes sistemas se tornassem um tema de pesquisa relevante com aplicações de interesse das comunidades científica e civil. A resolução dos sistemas SAR é da ordem do comprimento de onda do sinal transmitido e por apresentar a frequência de operação na faixa de microondas os sinais possuem alta capacidade de penetrar locais com vegetação leve e detectar estruturas subterrâneas contanto que estejam enterradas em ambientes de solo seco (YUE et al., 2021).

Devido a estas características há uma gama de aplicações que utilizam imagens SAR, tal como detecção de veículos na vegetação, monitoramento de áreas florestais, medição topográfica e avaliação de desastres naturais (YUE et al., 2021). Considera-se a distribuição Rayleigh para modelar a amplitude dos registros desse radar.

O sistema SAR emite radiação e realiza o registro dos sinais de retorno em uma área denominada resolução de célula. Um sinal de retorno pode ser interpretado como um somatório dos sinais refletidos por vários objetos situados na célula, também denominados espalhadores. Existem diferentes modelagens estatísticas a partir do sinal refletido e do número desses espalhadores. O modelo Rayleigh assume que o número de espalhadores equivalentes por resolução de célula de cada imagem satisfaz o teorema do limite central. Contudo, é comum que os sinais de retorno sejam originados a partir de um número reduzido de espalhadores o que faz com que esta suposição não seja satisfeita.

1.1 OBJETIVO

Este trabalho de conclusão de curso visa estudar, a partir da onda eletromagnética retroespalhada por objetos, a modelagem estatística de áreas com diferentes características de superfície, especificamente: veículos na floresta, floresta, lago e estrada. Além disso, propõe investigar a relação destas regiões com o modelo Rayleigh, ou modelo do Speckle completamente desenvolvido, com base em uma abordagem quantitativa para o cálculo do número de elementos espalhadores, a partir da análise do Speckle. Isto é, da degradação granular observada em imagens devido as variações de intensidade entre os pixels. Tal degradação é causada pela interferência ou adição de sinais refletidos por vários espalhadores, combinados de forma aleatória em um sinal retro difundido.

Utilizou-se a metodologia desenvolvida por (XU, 2022) para investigar a relação do número de espalhadores entre as regiões citadas anteriormente e as distribuições de melhor ajuste aos dados. A partir disso, espera-se que o presente trabalho contribua para

a maior compreensão de imagens SAR e que possa servir como referência a trabalhos futuros principalmente aplicações de detecção de alvos usando imagens SAR.

1.2 ORGANIZAÇÃO DO TRABALHO

Este trabalho é dividido em sete capítulos, os quais tratam:

1. Introdução: Contextualização e apresentação do tema que consiste no estudo da modelagem estatística de imagens SAR.
2. Fundamentação teórica: Introdução ao sistema de imageamento SAR pertinente ao entendimento do estudo realizado.
3. Modelagem Estatística: Introdução as abordagens de modelagem estatística de imagens SAR.
4. Materiais e métodos: Metodologia e conjunto de dados usados na simulação.
5. Resultados: Resultados das simulações contendo a modelagem estatística e cálculo do número de espalhadores para as regiões do estudo.
6. Análise dos resultados: Análise das simulações do capítulo anterior e comparação com os resultados do estudo de referência.
7. Conclusão: Retomada dos principais resultados e sugestão de possíveis trabalhos.

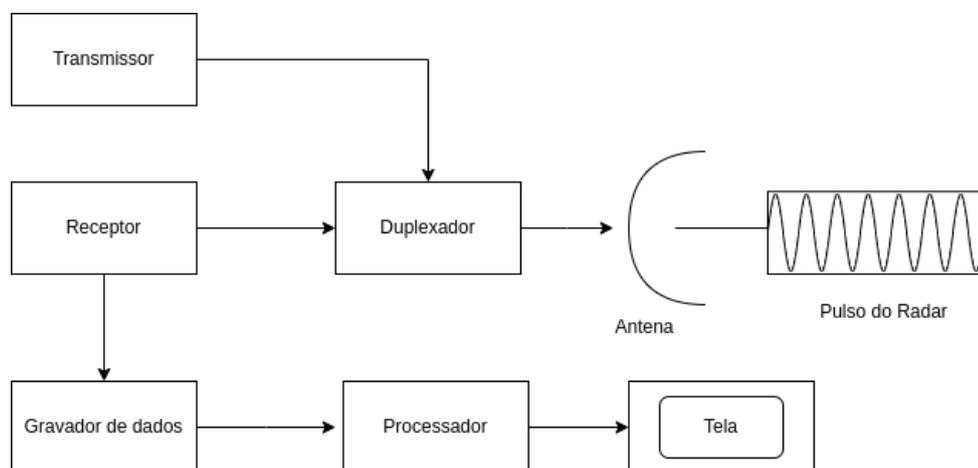
2 FUNDAMENTAÇÃO TEÓRICA

2.1 RADAR DE ABERTURA SINTÉTICA

De acordo com (TREVETT, 2008) o radar é um dispositivo que é capaz de detectar um objeto (alvo), indicando sua distância (*range*) e posição (direção). Este sistema de detecção pode ser usado como uma tecnologia de imageamento ativo, no qual pulsos de energia na faixa de microondas são emitidos por uma antena e as reflexões resultantes são utilizadas para formar imagens (AGER, 2013),

Uma configuração típica de radar com transmissor, duplexador, antena, receptor e gravador de dados é ilustrada na Figura 2.1. O transmissor gera ondas eletromagnéticas de alta energia na frequência de rádio e o duplexador direciona o pulso para a antena e retorna o eco para o receptor. A antena transmite pulsos eletromagnéticos em direção a área a ser imageada e coleta os ecos. O sinal retornado é convertido em um número digital pelo receptor e o gravador de dados armazena os valores para posterior processamento e exibição (KOO, 2008).

Figura 2.1 – Diagrama de blocos dos componentes básicos de um sistema de radar.



Fonte: Adaptado de Koo (2008)

O sistema SAR é um radar de imageamento acoplado a uma plataforma móvel com ampla utilização no sensoriamento remoto por possibilitar imagens de alta resolução independente das condições atmosféricas. De forma similar aos radares convencionais, ondas eletromagnéticas são sequencialmente transmitidas e os ecos são captados pela antena do radar (MOREIRA et al., 2013). O atributo básico deste sistema de imageamento é a abertura sintética, que consiste na combinação dos sinais refletidos ao longo de uma trilha de voo para sintetizar uma grande antena (MEDEIROS, 1999).

A Tabela 2.1 apresenta as bandas de frequência comumente adotadas nos sistemas SAR com os frequências e comprimento de ondas associados, em que uma maior penetração dos pulsos eletromagnéticos ocorre para comprimentos de ondas mais longos (INPE, 2019).

Tabela 2.1 – Bandas de frequência, faixas de comprimento de onda e frequências associadas comumente usadas em sistemas SAR.

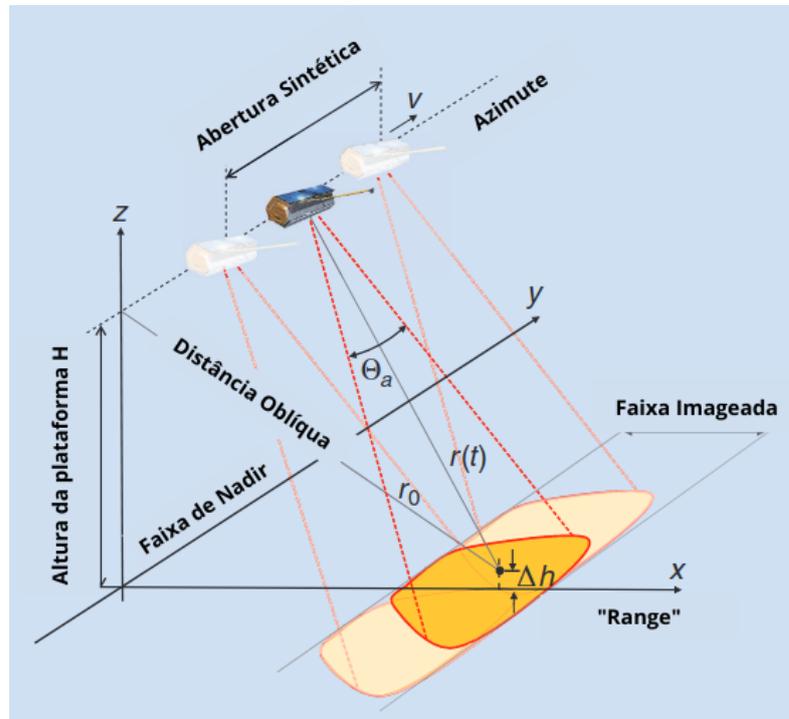
Banda de Frequência	Comprimento de Onda (cm)	Frequência (GHz)
KA	0,75-1,2	40-25
Ku	1,7-2,5	17,6-12
X	2,5-4	12-7,5
C	4-8	7,5-3,75
S	8-15	3,75-2
L	15-30	2-1
P	60-120	0,5-0,25

Fonte: Adaptado de Moreira et al. (2013)

2.2 IMAGEAMENTO SAR

A geometria básica de um sistema de imageamento SAR é ilustrada na Figura 2.2. A plataforma com um sensor SAR é deslocada a uma velocidade V e altura H em relação ao solo. À medida que a plataforma se desloca o transmissor envia pulsos à uma taxa de repetição de frequência de pulso (PRF, do inglês *pulse repetition frequency*) e registra os ecos correspondentes para posterior processamento e formação da imagem (MEDEIROS, 1999).

Figura 2.2 – Geometria de imageamento SAR. r_o é a menor distância de aproximação, θ_a é a largura do feixe do azimute e v é a velocidade do sensor fixado na plataforma.



Fonte: Adaptado de Moreira et al. (2013)

As coordenadas da linha de imageamento do radar são medidas pela direção de vôo da plataforma conhecida como azimute e a direção de alcance (*range*). A antena aponta lateralmente formando um ângulo de visada θ em relação ao nadir. A faixa imageada mostrada na Figura 2.2 denota o alcance de imageamento do radar e seu comprimento depende da duração de captação dos dados (MOREIRA et al., 2013).

Segundo (MOREIRA et al., 2013) os sensores SAR geralmente utilizam sinais modulados em frequência (*chirp*), no qual a frequência varia linearmente com o tempo conforme a equação $f_i = k_\tau t$, onde k_τ é a taxa de *chirp*. O tempo na direção de varredura é denotado como *fast time* devido as ondas eletromagnéticas viajarem na velocidade da luz. A distância entre o radar a uma velocidade constante v e um ponto no solo descrito pelas coordenadas $(x, y, z) = (x_o, 0, \Delta_h)$ é obtida aplicando o teorema de Pitágoras.

$$r(t) = \sqrt{r_o^2 + (vt)^2} \approx r_o + \frac{(vt)^2}{2r_o} \text{ para } vt/r_o \ll 1 \quad (2.1)$$

em que $t = t_o = 0$ é o tempo de maior aproximação quando a distância é mínima e $r(t_o) = r_o = \sqrt{(H - \Delta_h)^2 + x_o^2}$ com a altura da plataforma H .

O tempo nesta expressão está associado com o movimento da plataforma e é denominado *slow time*. A variação na distância de um ponto no tempo está associado a fase azimute por $\phi(t) = -4\pi(t)/\lambda$, onde λ é o comprimento de onda do sinal.

O sistema SAR apresenta resolução em duas dimensões, a expressão para a reso-

lução *slant range* é $\delta_r = c_o/2B_r$, Onde c_o é a velocidade da luz e B_r é a largura de banda do sinal. Nota-se que esta resolução consiste na mínima distância em que dois pontos podem ser distinguidos (LU, 2019) e depende da largura de banda, que para sinais *chirp* corresponde a variação de frequência em um pulso. Portanto, uma alta resolução *slant range* pode ser obtida apenas através da variação de frequência em um pulso (AGER, 2013).

A resolução azimute δ_a é fornecida na construção da abertura sintética, que compreende o comprimento usado para captar os sinais de retorno. A largura do feixe da antena de comprimento d_a é aproximado por $\theta_a = \lambda/d_a$. O comprimento da abertura sintética é dado por $L_{sa} = \theta_a \dot{r}_o = \lambda/2L_{sa}$. A expressão da resolução azimute, segundo (MOREIRA et al., 2013),

$$\delta_a = \frac{d_a}{2} \quad (2.2)$$

A Equação 2.2 demonstra que a resolução azimute depende da dimensão da antena do SAR, sendo possível obter uma alta resolução independente da distância do radar e do alvo imageado.

O sinal de eco recebido forma um matriz bidimensional de valores complexos em que cada amostra apresenta um valor de amplitude e de fase. A primeira dimensão corresponde a direção *range* ou *fast time*, enquanto a segunda é chamada azimute ou *slow-time*. Assim, no sistema de imageamento SAR os sinais recebidos da cena iluminada são amostrados tanto em *fast time (range)* quanto em *slow-time (azimute)* (MOREIRA et al., 2013).

A imagem SAR é obtida após a etapa de processamento de sinais que pode ser entendida como a operação de dois filtros casados ao longo da dimensão *range* e azimute. A primeira etapa consiste em comprimir o sinal *chirp* para um pulso curto realizando a multiplicação no domínio da frequência pelo complexo conjugado do espectro do sinal *chirp* transmitido.

Esse procedimento resulta em um *range* comprimido que contém a informação relativa a distância do radar e um ponto no solo. Na compressão na dimensão azimute o sinal é convoluído com sua função de referência, que compreende o complexo conjugado da resposta esperada de um alvo na superfície. Considerando um ponto na dimensão *range*, o sinal azimute pode ser modelado conforme a Equação 2.3.

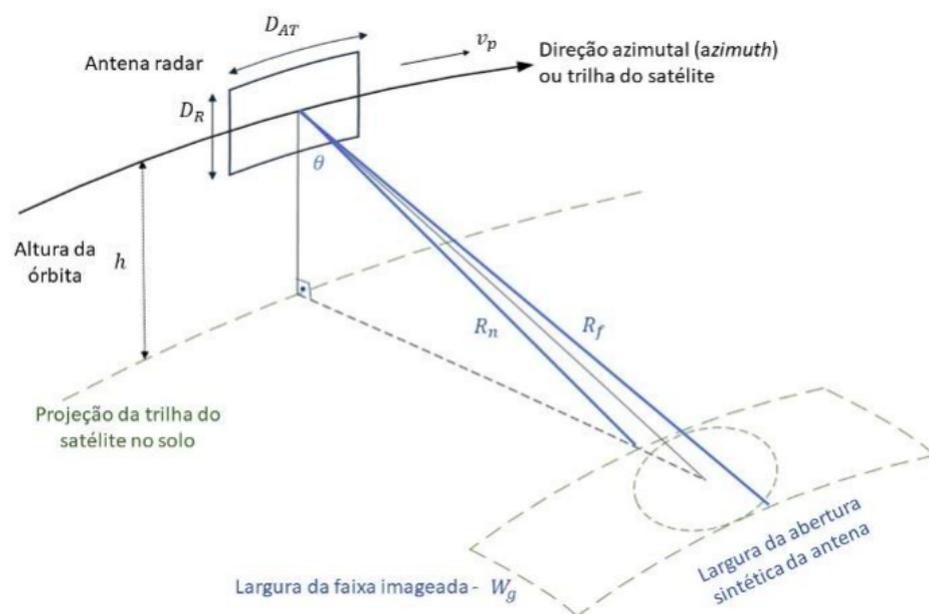
$$s_a(t) = A \sqrt{\sigma_0 \exp(i\phi^{scatt}) \exp(-i \frac{4\pi}{\lambda} r(t))} \quad (2.3)$$

onde A consiste nos parâmetros do sistema do sinal recebido tal como potência, perdas e a ponderação padrão da antena em função do azimute e ângulo de elevação; σ_0 é a seção transversal do radar (RCS, do inglês *radar cross section*); ϕ^{scatt} é a fase do sinal de retroespalhamento; $4\pi r(t)/\lambda$ descreve a variação da fase azimute devido a mudança na distância; e i é a unidade imaginária $i = \sqrt{-1}$.

2.3 TÉCNICAS DE IMAGEAMENTO SAR

Os principais modos de imageamento SAR são: *Stripmap*, *Spotlight* e *Scan*. No modo de operação *Stripmap* a antena do radar é direcionada no sentido perpendicular a trajetória com pulsos contínuos sendo enviados do início ao fim do imageamento uma faixa (OTE, 2019). Nesse modo de operação há uma maior resolução em azimuth, tendo em vista que são retornados mais sinais do que em um único *scan* (FRERY; WU; GOMEZ, 2022). A Figura 2.7 ilustra esse modo de aquisição.

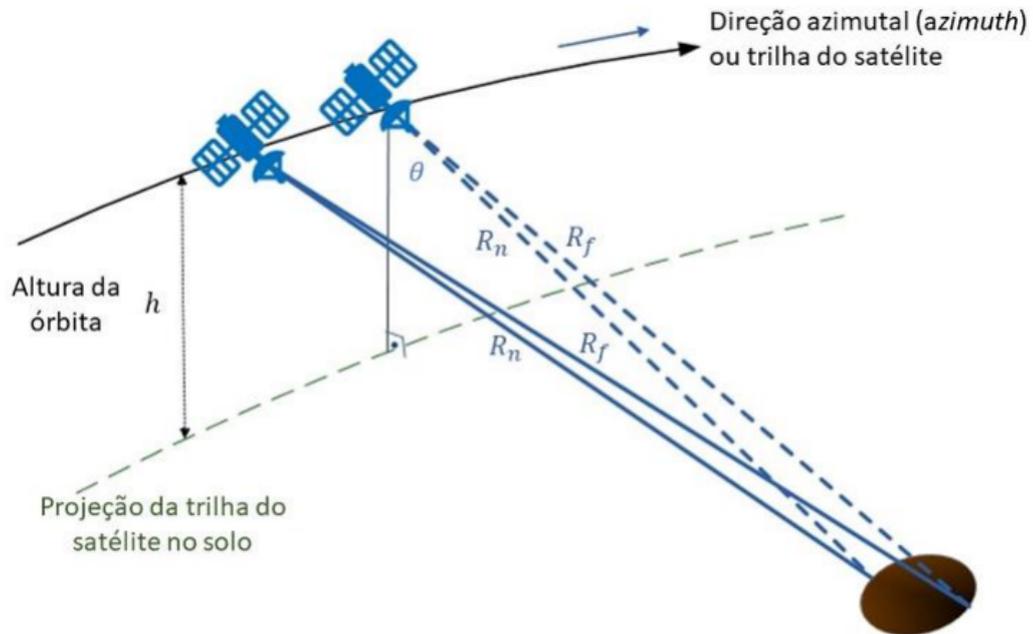
Figura 2.3 – Modo de aquisição de imagens Stripmap.



Fonte: Adaptado de OTE (2019)

No modo *Spotlight* a antena do radar é direcionada para uma região por um maior período de tempo (OTE, 2019) através do deslocamento do feixe do radar. A partir disso é possível obter uma maior resolução espacial do que no modo *Stripmap* ao custo de uma menor faixa imageada.

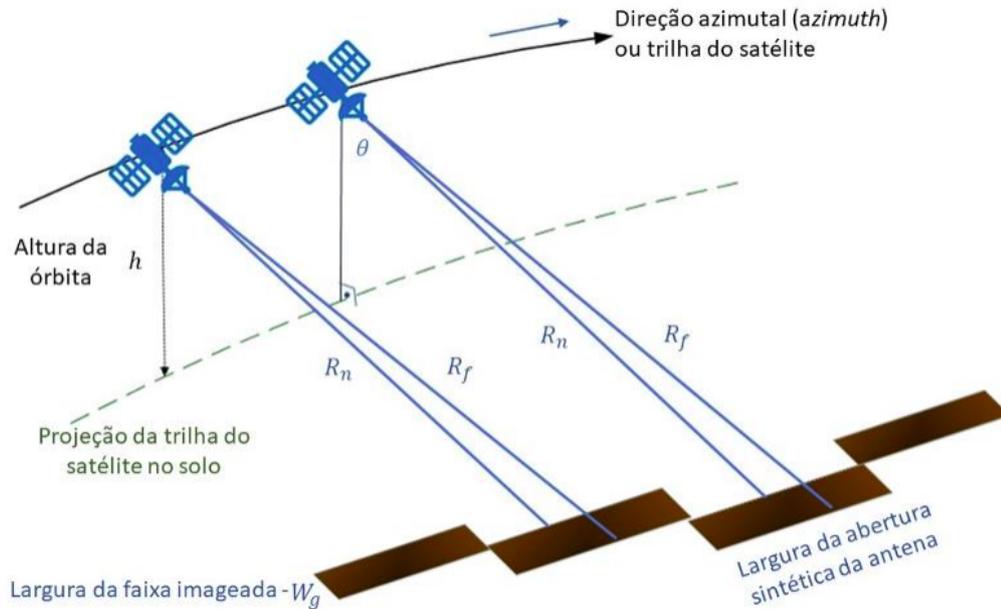
Figura 2.4 – Modo de aquisição de imagens Spotlight.



Fonte: Adaptado de OTE (2019)

O modo *SCAN* a antena do radar faz uma varredura periódica de uma região em diferentes orientações imageando subfaixas no solo, o que permite aumentar a faixa imageada. No entanto, há perdas na resolução espacial em função da redução de sinais retornados (OTE, 2019). Esse modo é interessante caso seja necessário imagear uma região de maior área sem precisar de uma alta resolução.

Figura 2.5 – Modo de aquisição de imagens Scan.



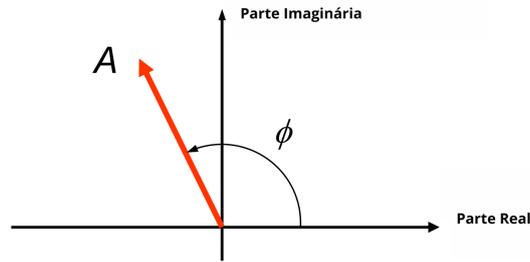
Fonte: Adaptado de OTE (2019)

2.4 TIPOS DE IMAGENS SAR

No imageamento SAR a estimativa da refletividade em cada pixel é representado por um número complexo que pode ser escrito na forma de fasor $Ae^{i\phi}$. São necessárias duas etapas para representação correta dos dados SAR nesse formato: calibração e geocodificação. A primeira garante que o valor de cada pixel represente características de espalhamento da região imageada e a etapa de geocodificação assegura que a localização de cada pixel na imagem SAR esteja associada a uma posição na superfície (TREVETT, 2008; MOREIRA et al., 2013).

Outras representações podem ser obtidas a partir da imagem complexa SAR, no qual as mais comuns são as imagens no formato de amplitude A e intensidade $I = A^2$ ou \log intensidade. Apesar destas imagens apresentarem ruído, são mais simples de serem interpretadas sendo possível a detecção visual de estruturas. A Figura 2.6 mostra a representação do pixel na imagem SAR com sua amplitude, associada ao coeficiente de retroespalhamento (σ^0), e fase (ϕ).

Figura 2.6 – Representação do pixel no formato complexo.



Fonte: Adaptado de Marwan (2015)

2.5 SINAL RETROESPALHADO

A seção da área transversal do radar (RCS) mede a capacidade de detecção de um objeto que é refletido no radar, parte da energia transmitida pelo radar ao "iluminar" uma região. O RCS é conhecido como uma assinatura eletromagnética de um objeto e é denotado por σ (FRERY; WU; GOMEZ, 2022), alguns fatores que influenciam o seu valor são: o tamanho do alvo em relação ao comprimento de onda do sinal do radar, a composição do alvo, o ângulo de incidência do sinal emitido pelo radar e a polarização do sinal.

O sinal refletido ou retornado pelo radar depende de fatores como a característica física do objeto imageado e as propriedades do sinal transmitido pelo radar (FRERY; WU; GOMEZ, 2022). Em regiões de floresta geralmente o sinal transmitido cria uma reflexão difusa com os sinais retornados de intensidade moderada e pixel cinzas na imagem. No entanto, destaca-se que caso o comprimento de onda do sinal seja de maior comprimento - como da banda UHF - eles não iriam ser refletidos nas árvores, mas sim na superfície e nos troncos das árvores com objetos na floresta podendo ser visualizados (AGER, 2013).

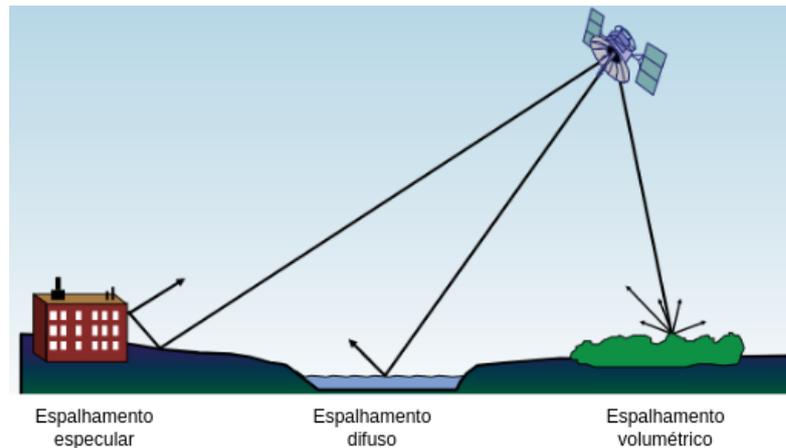
A rugosidade é um dos fatores que interferem na resposta da superfície e é uma grandeza relativa que depende do comprimento de onda (CENTENO, 2019). A superfície pode ser lisa ou ter aparência rugosa dependendo do comprimento de onda. Os sinais de retorno de menor rugosidade tem pouca energia retornada ao radar gerando imagens em tons escuros. Caso a superfície da água seja mais rugosa, como no caso do mar aberto, as reflexões das ondas geralmente são relativamente mais claras comparadas com o fundo (AGER, 2013). Em regiões urbanas de superfície mais rugosa, apresentam um maior retorno do sinal também gerando regiões mais claras pela maior amplitude do sinal retornado. O retroespalhamento pode ser definido em três tipos (CENTENO, 2019),

- Difuso: Apenas uma pequena parte do sinal emitido é retornado em direção à antena

- De volume: O sinal é refletido por várias superfícies pela superposição de várias camadas.
- Especular: Ocorre a reflexão total do sinal.

A Figura 2.7 ilustra os três tipos de espalhamento mencionados anteriormente.

Figura 2.7 – Variações de retroespalhamento.



Fonte: Adaptado de Ager (2013)

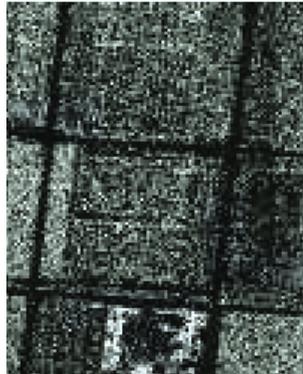
2.6 SPECKLE

A degradação granular chamada *speckle* é comum em todos os processos de formação de imagem com iluminação coerente tal como o empregado pelo radar de abertura sintética. Este fenômeno surge devido as variações de intensidade entre os pixels de uma imagem causadas pela interferência ou adição de múltiplos dispersores combinados de forma aleatória em um sinal retro difundido.

Em imagens SAR o *speckle* dificulta sua interpretação e análise e reduz a eficácia de aplicações como segmentação de imagens e classificação de objetos. A Figura 2.8 demonstra o aspecto de ruído com padrão granular *speckle* em uma imagem SAR (OLIVER; QUEGAN, 2004; TREVETT, 2008; LEE; POTTIER, 2017).

O método de operação do radar de abertura sintética pode gerar um número de observações ou visadas por pixel denotadas por L , Imagens com $L > 1$ são chamadas de imagens de múltipla visada. Durante o processamento da imagem a redução do efeito do *speckle* é feita pela média da intensidade da imagem de cada visada. A desvantagem desta técnica é que há perda de resolução por um fator L , o que reduz a informação de intensidade do pixel (OLIVER; QUEGAN, 2004; LEE; POTTIER, 2017).

Figura 2.8 – Imagem SAR com resolução 1m x 1m na banda de frequência Ku com as variações pixel a pixel no tom da imagem características do speckle.



Fonte: Adaptado de Ulaby et al. (2014)

2.7 ESTATÍSTICA DESCRITIVA

A estatística descritiva é útil ao selecionar um modelo ou método para análise de dados (FRERY; WU; GOMEZ, 2022). Segundo (VIEIRA, 2018) a estatística descritiva consiste nos procedimentos usados para organizar a coleta, apuração e descrição dos dados. Algumas propriedades que servem como base para a análise dos resultados e das distribuições trabalhadas neste estudo serão apresentadas nesta seção.

2.7.1 Medidas de tendência central

As medidas de média, mediana e a moda são quantidades que podem ser usadas para descrever a tendência central de um conjunto de dados (FRERY; WU; GOMEZ, 2022). A média é definida como,

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \quad (2.4)$$

onde n é o conjunto de amostras de um conjunto de dados e x_i é uma amostra. A média pode não ser a melhor medida para descrever um conjunto de dados com valores extremos.

A mediana amostral é obtida pela ordenação das n observações da menor para a maior (com os valores repetidos incluídos, de forma que cada observação da amostra seja exibida na lista ordenada) (DEVORE; SILVA, 2006). Quando há valores extremos no conjunto de dados essa medida é mais adequada para descrever os dados. Ela pode ser

definida como,

$$M_e = \begin{cases} x_{\frac{n+1}{2}} & \text{se } n \text{ é ímpar} \\ \frac{x_{n/2} + x_{n/2+1}}{2} & \text{se } n \text{ é par} \end{cases} \quad (2.5)$$

A mediana amostral é obtida pela ordenação das n observações da menor para a maior (com os valores repetidos incluídos, de forma que cada observação da amostra seja exibida na lista ordenada)

2.7.2 Medidas de dispersão

As medidas de dispersão podem ser definidas como o afastamento de todos os valores de um conjunto de dados em relação a média aritmética ou mediana (UFS, 2019). Nessa subseção serão apresentadas algumas das principais medidas de dispersão.

A variância é a média quadrática dos desvios em relação a média aritmética, no caso de um conjunto de amostras a variância pode ser definida como,

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1} \quad (2.6)$$

A medida mais apropriada para medir o espalhamento dos dados é o desvio padrão por estar na mesma unidade, ela é definida como $s = \sqrt{s^2}$ (FRERY; WU; GOMEZ, 2022). Outra medida que expressa a variabilidade em relação a média é o coeficiente de variação que é um valor adimensional definido a razão entre o desvio padrão das amostras pela média,

$$CoV = \frac{s}{\bar{x}} \quad (2.7)$$

O quartis dividem os dados em quatro partes de modo que cada porção contenha 25% dos dados. Uma medida bastante usada para medir a variabilidade de uma distribuição é intervalo interquartil (IQR) definido como,

$$IQR = Q_3 - Q_1 \quad (2.8)$$

em que Q_3 e Q_1 representam 75% e 25% dos dados respectivamente.

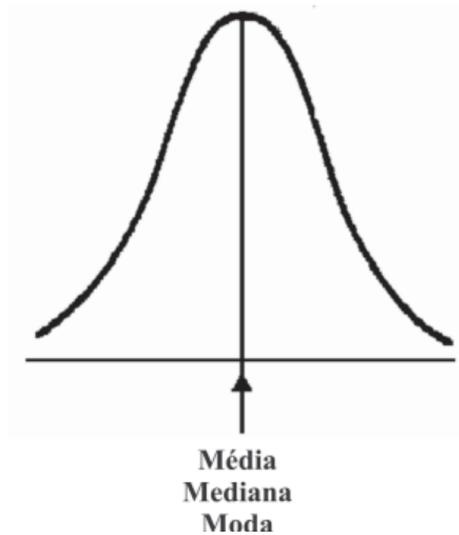
2.7.3 Formato dos dados

As medidas de assimetria e *Curtose* são utilizadas para descrever o formato de uma distribuição de dados. A assimetria S_k é definida como,

$$S_k = \frac{n \sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)(n-2)s^3} \quad (2.9)$$

A distribuição é considerada simétrica quando $S_k = 0$, quando a distribuição é simétrica a média, mediana e moda coincidem conforme mostra a Figura 2.9.

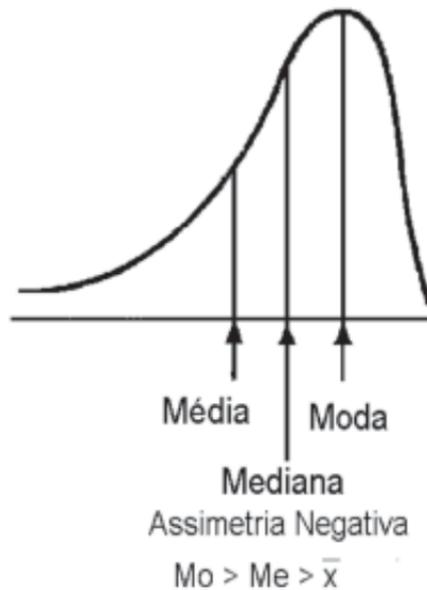
Figura 2.9 – Exemplo distribuição simétrica.



Fonte: Adaptado de UFS (2019)

Quando $S_k < 0$ a assimetria é considerada negativa ou enviesada a esquerda, já que existe uma maior concentração de valores na extremidade superior (UFS, 2019).

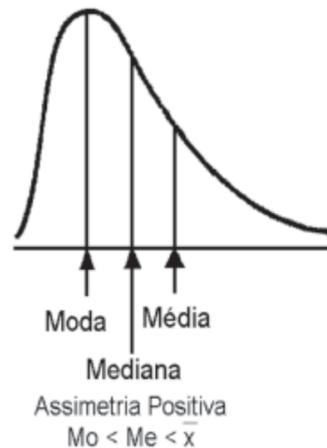
Figura 2.10 – Exemplo distribuição com assimetria negativa.



Fonte: Adaptado de UFS (2019)

A assimetria é positiva quando $S_k > 0$, nesse caso existe uma maior concentração na extremidade inferior da escala conforme mostra a Figura 2.11.

Figura 2.11 – Exemplo distribuição com assimetria positiva.



Fonte: Adaptado de UFS (2019)

A Curtose pode ser definida como o grau de achatamento de uma distribuição em relação a distribuição normal (UFS, 2019). Essa medida pode ser definida por,

$$K_u = \frac{n(n+1) \sum_{i=1}^n (x_i - \bar{x})^4 - 3(n-1)^3 s^4}{(n-1)(n-2)(n-3)s^4} \quad (2.10)$$

Quando a distribuição se aproxima de uma curva normal o valor de K_u irá ficar

próximo de zero; $K_u < 0$ significa que a curva é mais "achatada" que a distribuição normal e $K_u > 0$ representa uma distribuição que é mais estreita.

3 MODELAGEM ESTATÍSTICA DE IMAGENS SAR

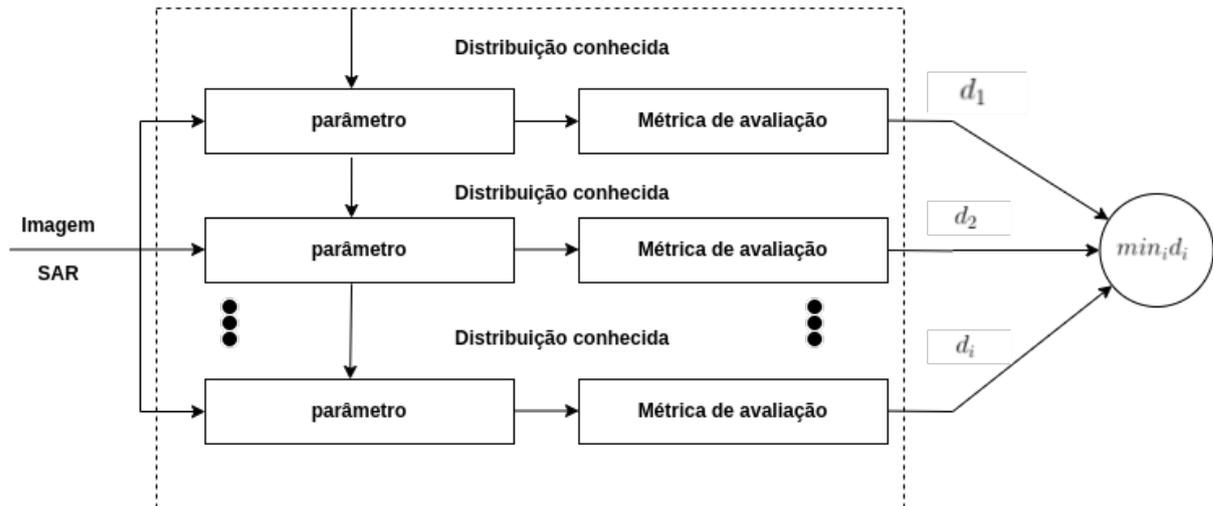
3.1 INTRODUÇÃO

A modelagem estatística de imagens é um dos fundamentos teóricos para interpretação de imagens SAR e através dela é possível obter uma compreensão do mecanismo de espalhamento da região imageada. Esta modelagem está presente em várias aplicações tais como supressão do speckle, classificação e detecção de alvos (GAO, 2010).

De acordo com (GAO, 2010) os modelos estatísticos de imagens SAR podem ser agrupados em duas categorias: modelos paramétricos e não paramétricos. Nos modelos paramétricos um conjunto de distribuições é analisada e os parâmetros da distribuição são desconhecidos e precisam ser estimados. Em modelos não paramétricos a distribuição de uma imagem não é assumida e o modelo que melhor descreve os dados são obtidos em uma abordagem orientada aos dados da imagem. Embora os modelos não paramétricos sejam mais flexíveis e consigam descrever os dados com mais acurácia, geralmente demandam um maior número de dados e um maior custo computacional (GAO, 2010).

O processo de modelagem paramétrico consiste nas etapas: (i) análise de modelos de distribuição estatísticos; (ii) estimação de parâmetros da distribuição; (iii) teste de qualidade de ajuste para verificar a acurácia dos modelos no ajuste de dados reais. O diagrama da Figura 3.1 representa as etapas da modelagem estatística usando modelos paramétricos. Busca-se estimar os parâmetros de um conjunto de distribuições conhecidas e o modelo que melhor descreve a imagem é escolhido de acordo com o valor ótimo de uma métrica na imagem representado pela distância mínima $\min_i d_i$.

Figura 3.1 – Diagrama de blocos do processo de modelagem estatística paramétrica.



Fonte: Adaptado de Gao (2010)

Nas seções seguintes são apresentados os modelos estatísticos adotados neste trabalho que descrevem o sinal de retorno em imagens SAR, estes pertencentes a categoria de modelos paramétricos.

3.2 MODELAGEM ESTATÍSTICA BASEADA NO MODELO DO ESPALHADOR COERENTE

O campo elétrico no receptor do radar dado um sinal de retorno devido ao elemento espalhador i pode ser matematicamente descrito como,

$$E_i = K_i E_i^0 e^{j\phi_i}, \quad \phi_i = \theta_i - 2kR_i \quad (3.1)$$

em que θ_i é o ângulo da fase; K_i é a constante do sistema; $k = 2\pi/\lambda$ é o número de onda; R_i é a distância do elemento espalhador até a antena do radar e E_i^0 é a intensidade do elemento espalhador (XU, 2022).

De acordo com (YUE et al., 2021), a primeira suposição do modelo do espalhador coerente é de que os espalhadores em uma resolução de célula são estatisticamente independentes. Assim, o campo elétrico E pode ser expresso como uma superposição coerente,

$$E = \sum_{i=1}^N K_i E_i^0 e^{j\phi_i} \quad (3.2)$$

A segunda suposição é de que a distância máxima entre os alvos $\Delta_r = |r_i - r_j|_{max}$ é muito menor que a distância média entre a antena e os alvos. Desse modo, o ganho da antena é uniforme entre toda a área iluminada $K_i = K$. Para simplificação, é usado $K = 1$,

assim,

$$E = \sum_{i=1}^N a_i e^{j\phi_i} = A e^{j\theta} \quad (3.3)$$

Em que A e θ são, respectivamente, a amplitude e fase do campo elétrico resultante E . A modelagem apresentada caracteriza o campo elétrico complexo E em um processo nomeado *random walk* no plano complexo que pode ser expresso como o somatório de N elementos espalhadores independentes em uma resolução de célula (OLIVER; QUEGAN, 2004).

Os modelos do speckle Rayleigh e NR (*Non-Rayleigh*) são derivados a partir desta modelagem e serão abordados a seguir.

3.2.1 Modelo Rayleigh do speckle

O modelo Rayleigh do speckle é adotado para descrever os sinais de retorno de uma região uniforme de imagens de radar de baixa resolução. De acordo com (YUE et al., 2021) este modelo é obtido quando o modelo do espalhador coerente satisfaz as seguintes condições:

1. As amplitudes e fases são estatisticamente independentes.
2. Devido a fase de cada elemento $\phi_i = \theta_i - 2kr_i$ ser uma variável aleatória determinada pela posição r_i , então a amplitude a_i e r_i são também independentes.
3. As amplitudes dos elementos espalhadores são independentes e identicamente distribuídas (iid) com primeiro e segundos momentos finitos, que garante que não há espalhadores dominantes em uma resolução de célula.
4. A localização dos elementos espalhadores é aleatória e a distância máxima Δ_r é muito maior que o comprimento de onda do radar de forma que a fase do elemento ϕ_i é distribuída de modo uniforme no intervalo $(-\pi, \pi)$.
5. Há um grande número de espalhadores N em uma resolução de célula.

Segundo (LEE; POTTIER, 2017) quando estas condições são atendidas elas descrevem o speckle completamente desenvolvido. As componentes reais e imaginárias do campo elétrico resultante E podem ser expressas como,

$$A_x = \text{Re}\{E\} = \sum_i^N a_i \cos\phi_i \quad (3.4)$$

$$A_y = \text{Im}\{E\} = \sum_i^N a_i \sin\phi_i \quad (3.5)$$

O teorema do limite central pode ser utilizado quando há um grande número de elementos espalhadores N em uma resolução de célula. Nesse caso, as componentes A_x e A_y podem ser tratadas como variáveis aleatórias independentes e normalmente distribuídas com média zero e variância denotada por $\sigma^2/2$ (LEE; POTTIER, 2017).

A função densidade da amplitude A definida como $A = \sqrt{A_x^2 + A_y^2}$ segue a distribuição Rayleigh para dados SAR capturados em uma visada (LEE; POTTIER, 2017),

$$p(A) = \frac{2A}{\sigma^2} e^{-\frac{A^2}{\sigma^2}} \quad (3.6)$$

A intensidade é definida como $I = A_x^2 + A_y^2 = A^2$ e segue uma distribuição exponencial negativa.

$$p(I) = \frac{1}{\sigma^2} e^{-\frac{I}{\sigma^2}} \quad (3.7)$$

Outras informações relevantes sobre esta distribuição podem ser encontradas em (LEE; POTTIER, 2017) e (YUE et al., 2021),

3.2.2 Modelagem estatística baseada no modelo speckle NR

De acordo com (YUE et al., 2021) o modelo do speckle NR (*Non-Rayleigh*) é baseado no modelo do espalhador coerente e considera que o sinal de retorno do radar contém um sinal alvo com variações devido as seguintes explicações:

1. A amplitude do elemento espalhador em uma resolução de célula não é constante e apresenta variações características.
2. O teorema do limite central não é satisfeito devido ao número de espalhadores N em uma resolução de célula ser pequeno.
3. A distribuição da fase do campo resultante em uma resolução de célula não é distribuída de modo uniforme.

Dessa forma, devido ao número insuficiente de espalhadores N para satisfazer o teorema do limite central o modelo Rayleigh não pode ser adotado. A seguir é apresentada a distribuição K e a distribuição G para modelagem do speckle NR.

3.2.2.1 Distribuição K

A distribuição K é baseada no processo físico do espalhamento (LEE; POTTIER, 2017) e geralmente garante uma boa qualidade de ajuste em regiões heterogêneas em imagens de alta resolução (GAO, 2010).

De acordo com (LEE; POTTIER, 2017) a derivação desta distribuição é baseada no processo físico do espalhamento e sua modelagem estatística considera que a amplitude e intensidade dos elementos espalhadores seguem a distribuição K e que o número de espalhadores N é uma constante finita de valor arbitrário (YUE et al., 2021).

A amplitude da distribuição pode ser caracterizada como (XU, 2022),

$$f_k(A) = \frac{2}{a\Gamma[n(1+v)]} \left(\frac{A}{2a}\right)^{n(1+v)} K_{n(1+v)-1}\left(\frac{A}{a}\right) \quad (3.8)$$

onde $v > -1$ e a são parâmetros dependentes da cena, $K_n(\cdot)$ é uma função Bessel modificada de segundo tipo e $\Gamma(\cdot)$ é a função Gama.

3.2.2.2 Distribuição G^0

O modelo multiplicativo modela o sinal de retorno de uma imagem SAR como o produto $Z = XY$, com X sendo o retroespalhamento da superfície imageada e Y o *speckle*. O modelo G^0 considera o sinal de retorno como o produto de duas variáveis aleatórias $X \sim \Gamma^{-1}(\alpha, \gamma)$ e $Y \sim \Gamma(1, L)$ (FRERY; WU; GOMEZ, 2022). A função densidade é dada por,

$$f_Z(z; \alpha, \gamma, L) = \frac{L^L \Gamma(L - \alpha) Z^{L-1}}{\gamma^\alpha \Gamma(L) \Gamma(-\alpha) (\gamma + Lz)^{L-\alpha}} \quad (3.9)$$

onde $\alpha < 0$ é o parâmetro de rugosidade; $\gamma > 0$ é o parâmetro de escala; L é o número de visadas e $\Gamma(\cdot)$ é a função Gama.

Sendo que $Z \sim G^0(\alpha, \gamma, L)$ caracteriza o sinal de retorno seguindo a distribuição 3.9. Os momento de ordem k da distribuição são,

$$E(Z^k) = (\gamma/L)^k \frac{\Gamma(L+k)\Gamma(-\alpha-k)}{\Gamma(L)\Gamma(-\alpha)} \quad (3.10)$$

em que E representa o valor esperado. A distribuição G^0 tem o mesmo número de parâmetros da distribuição K , mas tem se mostrado mais apropriada para modelar dados com maior variabilidade como regiões urbanas (FRERY; WU; GOMEZ, 2022). Além disso, de acordo (FRERY; WU; GOMEZ, 2022) essa distribuição consegue modelar regiões similares as que a distribuição K propõe descrever.

3.3 MODELAGEM ESTATÍSTICA BASEADA NOS MODELOS EMPÍRICOS

A modelagem estatística baseada em modelos empíricos parte da análise da imagem e não considera o processo de espalhamento, tal como é feito no modelo do espalhador coerente. O objetivo principal desta modelagem é encontrar uma distribuição que é consistente com os dados reais de imagens SAR (YUE et al., 2021).

De acordo com (YUE et al., 2021) os modelos empíricos podem ser agrupados em três categorias: (i) distribuição única, (ii) mistura finita, e (iii) modelo não paramétrico. Neste estudo foram utilizadas distribuições da primeira categoria, que utilizam apenas uma distribuição matemática para modelar a imagem SAR.

3.3.1 Distribuição Lognormal

A distribuição lognormal é adequada para descrever dados de amplitude, superfícies urbanas e superfícies heterogêneas (MAÎTRE, 2013). Segundo (MARWAN, 2015) a função de densidade lognormal pode ser descrita como:

$$P(x) = \frac{1}{x\sqrt{2\pi}\sigma_L} e^{-\frac{(\log x - m_L)^2}{2\sigma_L^2}} \quad (3.11)$$

em que m_L e σ_L^2 são, respectivamente, a expectativa e a variância de $\log x$.

3.3.2 Distribuição Weibull

Segundo (GAO, 2010) a distribuição Weibull pode ser adotada para descrever imagens em uma única visada de forma satisfatória tanto em amplitude ou intensidade, porém não tem o mesmo desempenho em imagens com múltiplas visadas. A função densidade desta distribuição pode ser expressa como (YUE et al., 2021).

$$P_x(x) = \frac{x^n n}{\xi} \frac{1}{\xi} e^{-\frac{x}{\xi}} \quad (3.12)$$

onde ξ e n são, respectivamente, o parâmetro de escala e o parâmetro de forma da distribuição.

3.4 ESTIMAÇÃO DOS PARÂMETROS

Os métodos comumente usados na literatura para realizar a estimação de parâmetros de uma distribuição são o método dos momentos e da máxima verossimilhança (GAO,

2010). Neste trabalho, o método dos momentos foi usado para estimar os parâmetros da distribuição K conforme descrito em (XU, 2022), os parâmetros das distribuições Rayleigh, Lognormal e Weibull foram estimados usando o método da máxima verossimilhança.

3.4.1 Método dos momentos

Segundo (NICOLAS; ANFINSEN, 2002), assumindo que uma função densidade de probabilidade é descrita por p parâmetros, o método MoM (do inglês, *method of moments*) consiste em calcular os momentos das amostras de ordem 1 a p . Assim, um sistema de p equações em p parâmetros desconhecidos de funções de densidade de probabilidade é obtido. De forma geral, os k momentos dos dados $x_i, i = 1, \dots, N$, podem ser expressos como (XU, 2022),

$$m_k = \frac{1}{N} \sum_{i=1}^N x_i^k \quad (3.13)$$

A seguir é apresentado a estimação dos parâmetros a e n da distribuição K .

3.4.1.1 Estimação dos parâmetros da distribuição K

De acordo com (XU, 2022) o número de espalhadores da distribuição K , com base no método dos momentos (ISKANDER; ZOUBIR, 1999), pode ser estimado através da seguinte expressão,

$$n(1 + v) = \frac{2m_2^2}{m_4 - 2m_2^2} \quad (3.14)$$

em que v é o parâmetro de forma da distribuição K e m_i é o momento de ordem i dos dados analisados. O parâmetro de escala a pode ser descrito através da expressão por (XU, 2022),

$$a = \sqrt{\frac{m_3}{3m_1[n(1 + v) + 0,5]}} \quad (3.15)$$

3.4.2 Método da máxima verossimilhança

O método da Máxima Verossimilhança consiste em estimar os parâmetros de uma distribuição utilizando as estimativas que tornam máximo o valor da função de verossimilhança (BATISTA, 2009). De forma geral, o método pode ser caracterizada pela seguinte

expressão (NICOLAS; ANFINSEN, 2002),

$$\prod_{i=1}^N p_x(x_i) \quad (3.16)$$

onde \prod representa a operação de produto e $p_x(x_i)$ são os parâmetros que se deseja estimar para cada amostra x_i de um total de N amostras.

Ou de forma equivalente,

$$\sum_{i=1}^N \log(p_x(x_i)) \quad (3.17)$$

O ponto mínimo de uma função pode ser obtido a partir da solução da expressão da derivada parcial em relação a cada parâmetro da distribuição (BATISTA, 2009; NICOLAS; ANFINSEN, 2002),

$$\frac{\partial(\sum_{i=1}^N \log(p_x(x_i)))}{\partial \alpha_j} = 0 \quad (3.18)$$

Assim, este método depende da existência das derivadas parciais (NICOLAS; ANFINSEN, 2002).

3.4.3 Estimação dos parâmetros para cada distribuição

A Tabela 3.1 mostra os métodos de estimação utilizados para cada uma das distribuições usadas neste trabalho de conclusão, junto as funções de distribuição de cada distribuição e aos parâmetros que foram estimados.

Tabela 3.1 – PDFs das distribuições e métodos de estimação de parâmetros usados no trabalho

Distribuição	PDF	Parâmetros	Método de estimação
Rayleigh	$P(A) = \frac{2A}{\sigma^2} e^{-\frac{A^2}{\sigma^2}}$	σ^2	MLE
G0	$P_z(z; \alpha, \gamma, L) = \frac{L^\alpha \Gamma(L-\alpha)}{\gamma^\alpha \Gamma(L) \Gamma(-\alpha)} \frac{z^{L-1}}{(\gamma+Lz)^{L-\alpha}}$	α, γ	MLE
Weibull	$P(x) = \frac{x^n}{\xi} e^{-\frac{x^n}{\xi}}$	n, ξ	MLE
K	$P(A) = \frac{2}{a\Gamma[n(1+v)]} \left(\frac{A}{2a}\right)^{n(1+v)} K_{n(1+v)-1}\left(\frac{A}{a}\right)$	v, a, n	MoM
Lognormal	$P_x(x) = \frac{1}{x\sqrt{2\pi\sigma_L}} e^{-\frac{(\log x - m_L)^2}{2\sigma_L^2}}$	m_L, σ_L^2	MLE

3.5 RELAÇÃO DO NÚMERO DE ESPALHADORES E O MODELO DO SPECKLE

De acordo com (XU, 2022) em imagens SAR é difícil que o número de espalhadores presentes em uma resolução de célula seja elevado, principalmente em regiões homogêneas. Partindo do princípio que o número de espalhadores é grande o suficiente em seu trabalho (XU, 2022) demonstra que a partir de $N = 8$ o modelo do Speckle completamente desenvolvido passa a ser válido. Este valor é próximo ao encontrado em (ULABY et al., 2014) que descreve que a partir de $N = 10$ é suficiente para satisfazer o teorema do limite central e o modelo Rayleigh do Speckle passa a ser válido.

3.6 MÉTRICA DE AVALIAÇÃO

Nesta seção é apresentada a métrica adotada para avaliar a qualidade do ajuste das distribuições aos dados usados neste estudo. De acordo com (YUE et al., 2021) os principais métodos para avaliação de qualidade de um modelo são baseados em distância ou divergência. Nesse aspecto, há uma diversidade de testes que podem ser adotados que podem ser verificados em (YUE et al., 2021).

O teste de Kolmogorov-Smirnov foi utilizado neste trabalho para obter a qualidade dos modelos estatísticos usados para descrever os dados de imagens SAR. A sua escolha se deve a ser esta a métrica utilizada em (XU, 2022), estudo em que este trabalho é baseado e visa reproduzir. Utilizando a mesma métrica foi possível comparar os resultados deste estudo com os encontrados em (XU, 2022). A seguir é apresentado uma breve descrição deste teste.

3.6.1 Teste de Kolmogorov-Smirnov

O teste de Kolmogorov-Smirnov é usado para decidir se uma amostra vem de uma população com uma distribuição específica (RICCI, 2005). Este teste é baseado na comparação de uma distribuição acumulada empírica (ECDF) e a teórica definida por,

$$F(x) = \int_{\alpha}^x f(y, \theta) dy \quad (3.19)$$

em que $f(y, \theta)$ é a PDF. Dado n pontos X_1, X_2, \dots, X_n a ECDF pode ser definida como,

$$F_n(X_i) = \frac{N(i)}{n} \quad (3.20)$$

$N(i)$ é o número de pontos menor que X_i . O teste estatístico usado é dado por,

$$D_n = \sup|F(x_i) - F_n(x_i)|, 1 \leq i \leq n \quad (3.21)$$

O teste faz a diferença absoluta do extremo superior dos valores da ECDF e da CDF teórica (RICCI, 2005). A hipótese da forma da distribuição é rejeitada caso o teste estatístico D_n possua um valor p menor que um dado nível de significância.

4 MATERIAIS E MÉTODOS

Este capítulo apresentará uma descrição da base de dados usadas nas simulações, o ambiente de programação e a metodologia adotada.

4.1 DESCRIÇÃO DA BASE DE DADOS

Nesta seção será brevemente apresentado o sistema de radar usado para captura do conjunto de dados seguido de uma descrição à respeito das imagens utilizadas.

4.1.1 CARABAS-II

O sistema de radar SAR UWB CARABAS-II foi projetado pela Agência Sueca de Pesquisa em Defesa (FOI) e transmite ondas de rádio HH polarizadas em uma faixa de frequência de 20-90 MHz (LUNDBERG et al., 2006).

Uma resolução na ordem de 2,5 m x 2,5 m é obtida ao utilizar a largura de banda total do sinal e um ângulo de integração próximo a 70° (LUNDBERG et al., 2006). A Tabela 4.1 ilustra algumas das características deste sistema.

Tabela 4.1 – Parâmetros usados no sistema de radar CARABAS-II,

Parâmetros do sistema	Valor
Banda de frequência	20 - 86 MHz
Frequências centrais	21,25 - 85 MHz
Frequência de repetição de pulso	5 kHz
Modulação de Pulso	FM não linear
Alcance máximo	26,4 km
Altitude nominal de vôo	3 - 9 km
Velocidade nominal de vôo	127 m/s
Ângulo de abertura	90°
Potência transmitida (pico)	500 W

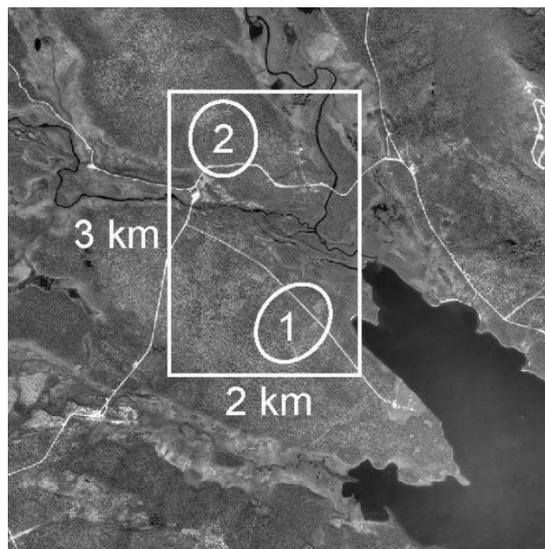
Fonte: Adaptado de Lundberg et al. (2006)

4.1.2 Conjunto de dados

O conjunto de dados consiste de 24 imagens SAR de amplitude registradas pelo sistema de radar SAR VHF UWB CARABAS-II durante uma campanha de voo em uma região militar no norte da Suécia chamada RFN Vidsele (LUNDBERG et al., 2006).

Conforme descrito em (LUNDBERG et al., 2006) as imagens deste conjunto de dados já estão calibradas, pré-processadas e georreferenciadas. O propósito da construção desta base de dados foi a coleta de dados SAR VHF para avaliar a performance de algoritmos de detecção de mudança em alvos em florestas. As regiões capturadas pelo sensor do sistema de radar compreendem áreas de florestas, lagos, campos e prédios. As imagens desta base de dados cobrem uma área de 6km^2 (2 km x 3km) conforme indicado pelo retângulo branco da Figura 4.1. O tamanho do pixel é de 1m x 1m, assim, a imagem forma uma matriz de dimensões 3000 x 2000 (LUNDBERG et al., 2006). As regiões em branco numeradas como 1 e 2 correspondem as localizações alvos na floresta dos pontos do terreno (deployments).

Figura 4.1 – Imagem aérea da região de teste dentro da localidade RFN Vidsele.

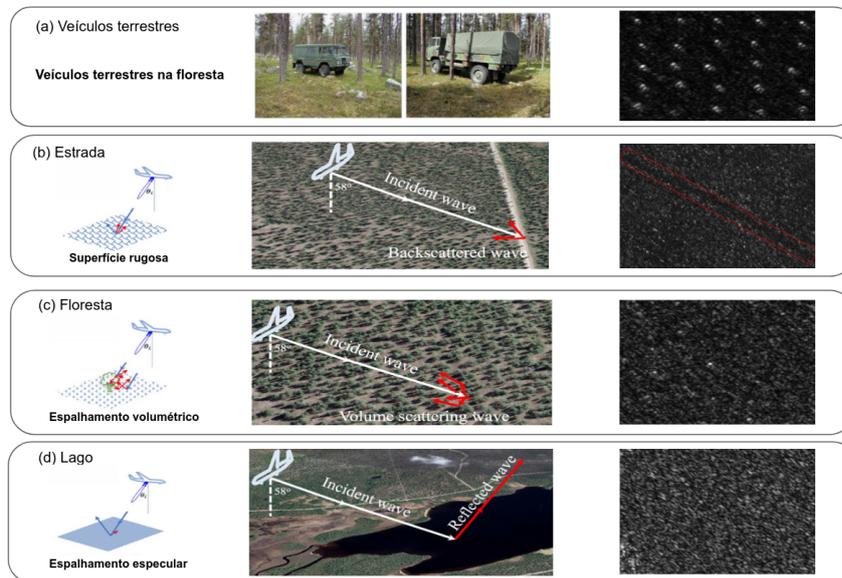


Fonte: Adaptado de Lundberg et al. (2006)

As imagens SAR foram selecionadas de um total de quatro missões de voo separadas com uma parcela de seis imagens de cada operação. Cada missão apresenta 25 alvos dispostos em diferentes *deployments* nomeados: Sigismund, Karl, Fredrik e Adolf-Fredrik. Os alvos consistem dos veículos militares: TGB11, TGB30, TGB40.

Em cada imagem foram selecionadas as regiões com os veículos na floresta, florestas, lago e estrada com o objetivo de analisar o número de espalhadores presentes e as distribuições de melhor ajuste para essas regiões mostradas na Figura 4.2. A seleção dos veículos na floresta foi feita com base na lista de coordenadas disponibilizadas junto ao conjunto de dados.

Figura 4.2 – Regiões com diferentes espalhamentos selecionadas a partir do conjunto de dados CARABAS-II para análise do processo de espalhamento e das distribuições com melhor ajuste. (a) Veículo Terrestre, (b) Estrada, (c) Floresta, (d) Lago.



Fonte: Adaptado de Lundberg et al. (2006), Xu (2022)

4.1.3 Metodologia

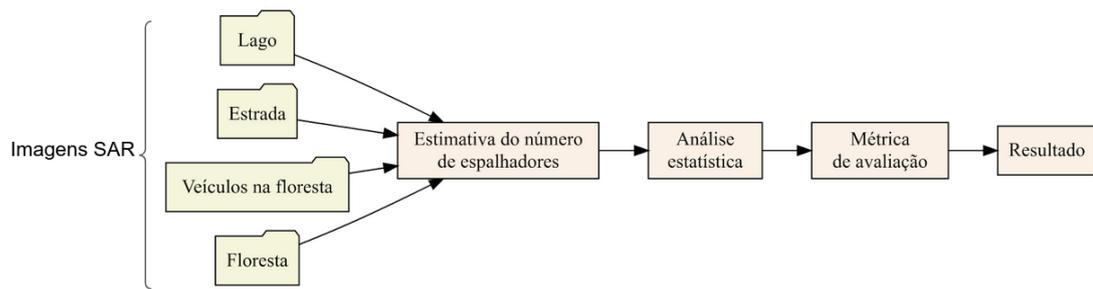
A metodologia para realização deste trabalho compreendeu as etapas de seleção dos dados, cálculo do número de espalhadores, análise estatística e comparação dos resultados com os apresentados no trabalho de (XU, 2022).

O método dos momentos apresentado na Seção 3.3 foi usado para estimar os parâmetros da distribuição K. A partir deste método foi obtido o número de espalhadores N para o conjunto de dados conforme a Equação 3.14. Os parâmetros das demais distribuições (G0, Weibull, Rayleigh, Lognormal) foram estimados usando MLE apresentada na Seção 3.4.

A métrica adotada para verificação da qualidade do ajuste das distribuições com as regiões comparadas foi a distância de Kolmogorov-Smirnov (KS) apresentada na Equação 3.19.

A Figura 4.3 ilustra o diagrama que representa as etapas usadas durante este trabalho. O pré-processamento compreendeu a seleção das imagens de cada região. Nesta etapa a linguagem de programação *Python* e as bibliotecas como *numpy* e *scipy* foram usadas na manipulação dos dados representadas no diagrama pelas imagens SAR. Nas demais etapas apresentadas no diagrama o software livre *R* foi usado.

Figura 4.3 – Diagrama de blocos das etapas desenvolvidas no trabalho.



Fonte: Próprio autor

5 RESULTADOS

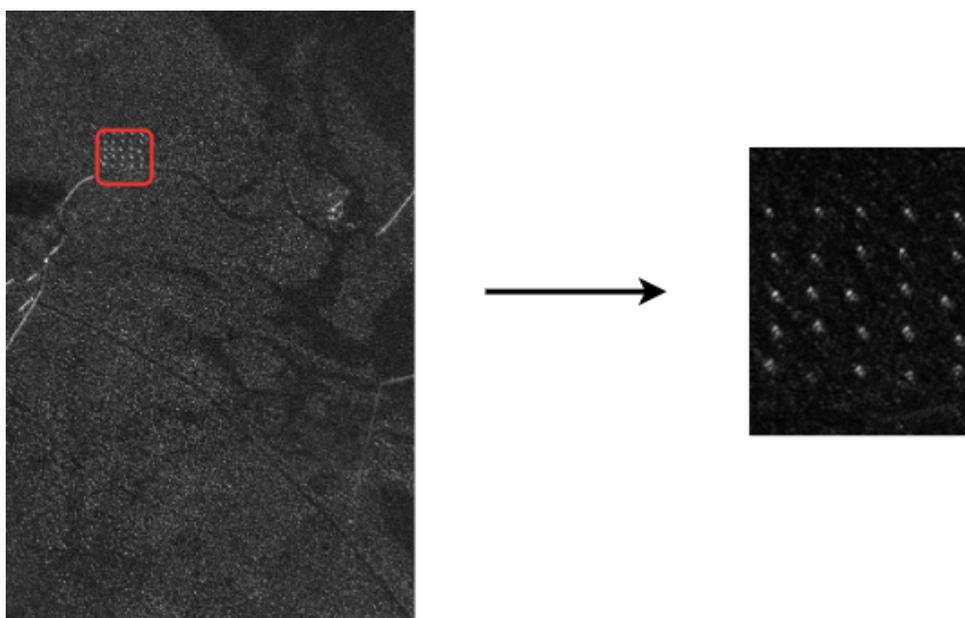
Este capítulo apresenta os resultados referentes ao número de espalhadores N estimados usando a Equação 3.14 e a qualidade de ajuste das distribuições para todas as 24 imagens da base de dados CARABAS-II. Os resultados foram agrupados por *deployment* e por região, a saber: veículos na floresta, floresta, estrada e lago.

São apresentados os momentos m_i usados para o cálculo do número de espalhadores usando a Equação 3.13. A distância KS foi usada para avaliar a qualidade de ajuste das distribuições aos dados de cada região, a Equação 3.19 foi usada para o cálculo desta métrica.

5.1 MODELAGEM ESTATÍSTICA DA REGIÃO DOS VEÍCULOS NA FLORESTA

A seleção das imagens da região dos veículos na floresta foi feita com base em uma lista com as coordenadas dos alvos de cada *deployment* conforme descrito na Seção 4.1.3. A Figura 5.1 apresenta a primeira imagem do *deployment* Sigismund com a região dos veículos selecionada para análise. Esta seleção foi feita para todas as imagens de cada *deployment*.

Figura 5.1 – Exemplo de imagem analisada da região dos veículos na floresta.



Fonte: Próprio autor

5.1.1 Deployment Sigismund

Tabela 5.1 – Número de espalhadores da região dos veículos do deployment Sigismund

Imagem	m_1	m_2	m_3	m_4	N
1	0,19	0,06	0,04	0,05	0,30
2	0,17	0,051	0,033	0,039	0,23
3	0,2	0,06	0,046	0,052	0,34
4	0,17	0,052	0,035	0,042	0,22
5	0,186	0,063	0,044	0,053	0,27
6	0,175	0,057	0,0402	0,0475	0,24

Tabela 5.2 – Distância KS das distribuições para a região dos veículos na floresta

Imagem	K	Rayleigh	LogNormal	Weibull	$G0$
1	0,324	0,189	0,049	0,072	0,162
2	0,336	0,205	0,0455	0,068	0,154
3	0,2961	0,174	0,048	0,068	0,164
4	0,3705	0,209	0,0476	0,0709	0,153
5	0,339	0,2116	0,047	0,0782	0,153
6	0,364	0,223	0,0468	0,065	0,158

5.1.2 Deployment Karl

Tabela 5.3 – Número de espalhadores da região dos veículos do deployment Karl

Imagem	m_1	m_2	m_3	m_4	N
1	0,183	0,0544	0,032	0,034	0,314
2	0,168	0,051	0,034	0,043	0,201
3	0,175	0,051	0,03	0,034	0,279
4	0,167	0,051	0,035	0,046	0,194
5	0,185	0,052	0,023	0,0154	0,826
6	0,157	0,044	0,027	0,031	0,213

Tabela 5.4 – Distância KS para a região dos veículos do deployment Karl

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,3099	0,145	0,053	0,06003	0,17
2	0,374	0,201	0,048	0,07	0,1574
3	0,333	0,154	0,053	0,062	0,1651
4	0,378	0,211	0,0478	0,072	0,154
5	0,218	0,12	0,0525	0,520	0,168
6	0,379	0,193	0,0497	0,007	0,1587

5.1.3 Deployment Fredrik

Tabela 5.5 – Número de espalhadores da região dos veículos Fredrik

Imagem	m_1	m_2	m_3	m_4	N
1	0,165	0,043	0,020	0,018	0,379
2	0,169	0,048	0,024	0,022	0,388
3	0,175	0,05	0,03	0,033	0,279
4	0,165	0,04	0,023	0,020	0,387
5	0,185	0,052	0,023	0,015	0,826
6	0,178	0,05	0,027	0,0284	0,324

Tabela 5.6 – Distância KS para a região dos veículos do deployment Fredrik

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,3	0,123	0,0535	0,051	0,171
2	0,352	0,173	0,0406	0,0056	0,153
3	0,379	0,154	0,053	0,0623	0,168
4	0,361	0,177	0,0398	0,057	0,152
5	0,2935	0,121	0,0520	0,0525	0,17
6	0,364	0,132	0,055	0,0563	0,169

5.1.4 Deployment Adolf-Fredrik

Tabela 5.7 – Número de espalhadores da região dos veículos do deployment Adolf-Fredrik

Imagem	m_1	m_2	m_3	m_4	N
1	0,165	0,04	0,015	0,0095	0,792
2	0,143	0,031	0,011	0,006	0,826
3	0,174	0,045	0,018	0,012	0,782
4	0,165	0,04	0,023	0,020	0,387
5	0,142	0,032	0,011	0,006	0,777

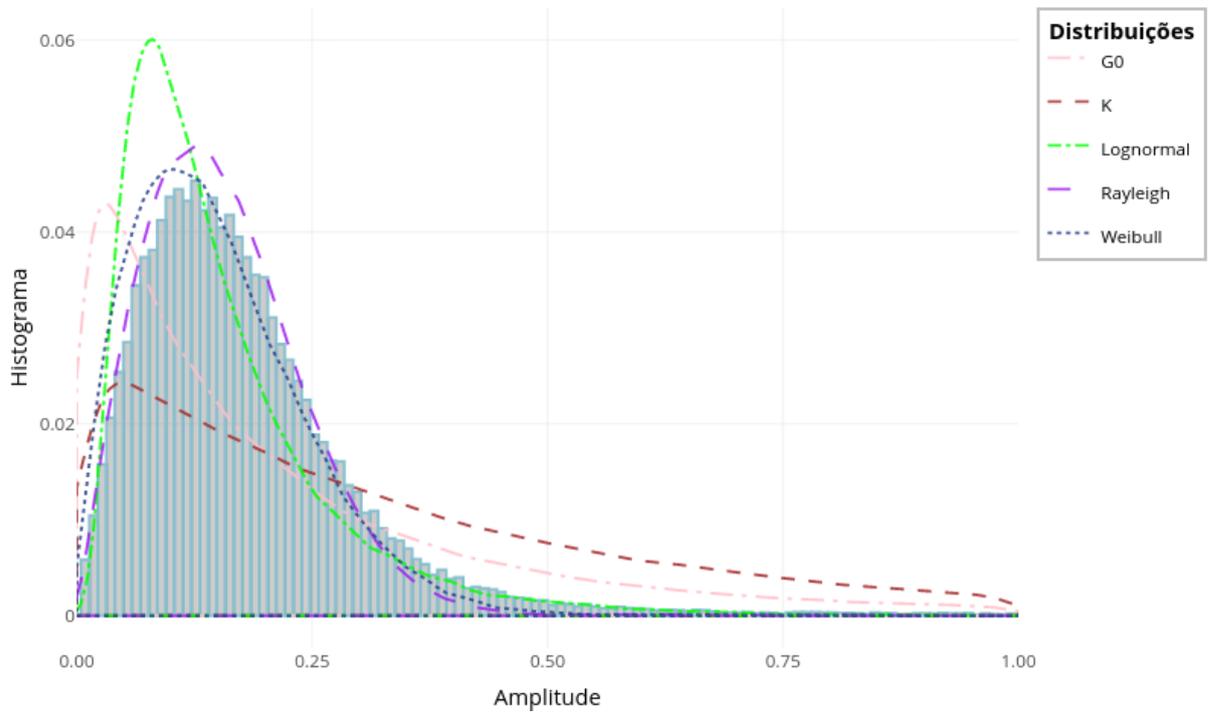
Tabela 5.8 – Distância KS para a região dos veículos do deployment Adolf-Fredrik

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,32	0,1054	0,0534	0,0437	0,171
2	0,348	0,143	0,0424	0,0475	0,161
3	0,305	0,1042	0,052	0,041	0,17
4	0,354	0,153	0,041	0,0515	0,157
5	0,297	0,090	0,055	0,041	0,179
6	0,365	0,132	0,055	0,0562	0,178

5.1.5 Visualização das distribuições da região dos veículos na floresta

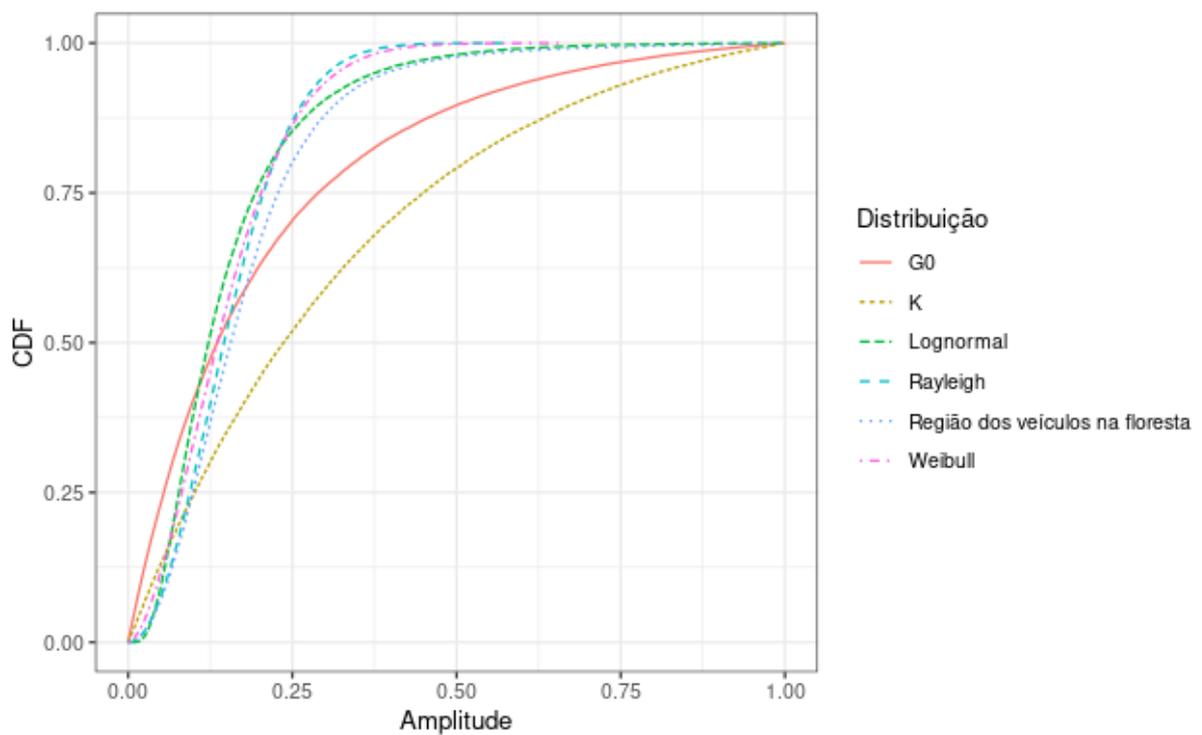
O histograma dos dados da primeira imagem do *deployment* Sigismund é mostrado na Figura 5.2 destacado em azul junto as funções de densidade de distribuição ajustadas aos dados. A função de distribuição acumulada empírica da região é mostrada na Figura 5.3.

Figura 5.2 – Histograma e PDF das distribuições da região dos veículos da imagem do *deployment* Sigismund.



Fonte: Próprio autor

Figura 5.3 – CDF da região dos veículos na floresta da imagem do deployment *sigismund*.



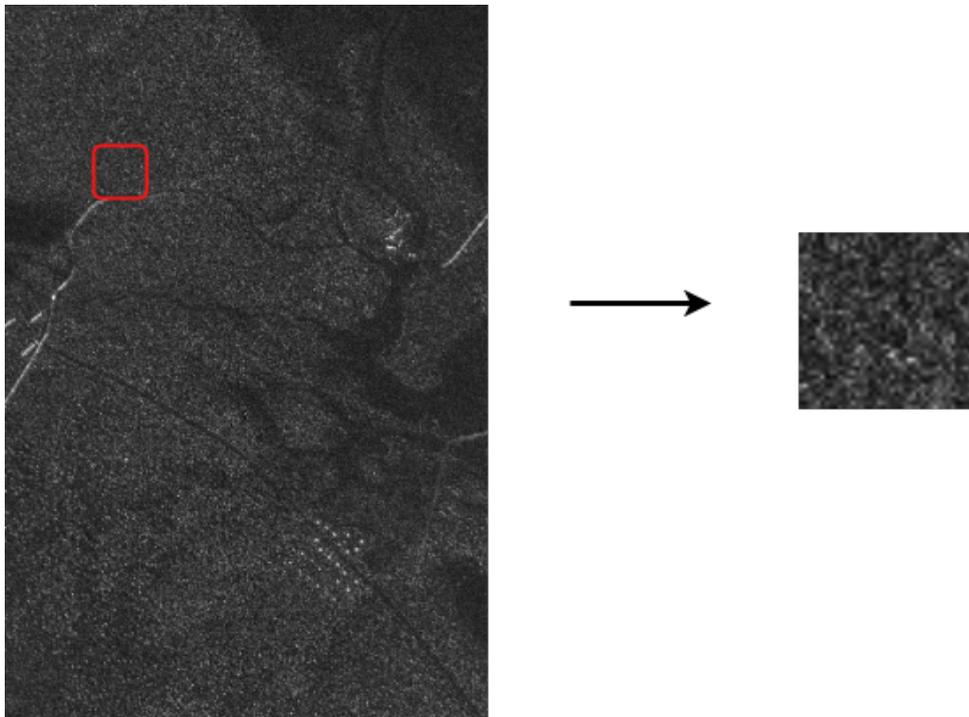
Fonte: Próprio autor

De acordo com o teste KS mostrado na Tabela 5.1, as distribuições que melhor descreveram os dados da região dos veículos na floresta foram as distribuições Lognormal. Este resultado pode ser visualizado de forma qualitativa na Figura 5.2 e na FDA da Figura 5.3.

5.2 MODELAGEM ESTATÍSTICA DA REGIÃO DA FLORESTA

A Figura 5.4 mostra o histograma da primeira imagem do *deployment* Fredrik com a região da floresta selecionada.

Figura 5.4 – Exemplo de imagem do *deployment* Fredrik para a análise da região da floresta.



Fonte: Próprio autor

5.2.1 Deployment Sigismund

Tabela 5.9 – Número de espalhadores da região da floresta Sigismund

Imagem	m_1	m_2	m_3	m_4	N
1	0,152	0,032	0,008	0,002	4,29
2	0,131	0,026	0,007	0,025	1,601
3	0,1605	0,35	0,0097	0,034	3,96
4	0,133	0,025	0,007	0,0025	1,75
5	0,166	0,037	0,01	0,0033	5,7
6	0,15	0,03	0,0076	0,002	4,17

Tabela 5.10 – Ajuste de distribuições usando a distância KS do estudo de referência

Imagem	K	Lognormal	Rayleigh	Weibull	G0
1	0,178	0,181	0,0517	0,02	0,18
2	0,125	0,043	0,123	0,0426	0,16
3	0,183	0,057	0,054	0,022	0,179
4	0,185	0,045	0,109	0,041	0,16
5	0,22	0,062	0,037	0,0169	0,189
6	0,18	0,059	0,0486	0,02	0,184

5.2.2 Deployment Karl

Tabela 5.11 – Número de espalhadores da região da floresta Karl

Imagem	m_1	m_2	m_3	m_4	N
1	0,155	0,033	0,008	0,003	4,007
2	0,137	0,027	0,0075	0,0027	1,85
3	0,149	0,003	0,0081	0,0027	3,392
4	0,134	0,0264	0,0728	0,027	1,615
5	0,174	0,0398	0,0111	0,003	11,17
6	0,154	0,0322	0,0086	0,0029	3,819

Tabela 5.12 – Distância KS para a região da floresta do deployment Karl

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,201	0,0527	0,0592	0,022	0,179
2	0,293	0,111	0,0463	0,0379	0,164
3	0,223	0,0629	0,0553	0,0264	0,18
4	0,313	0,114	0,0464	0,0397	0,16
5	0,191	0,0243	0,0645	0,0103	0,19
6	0,21	0,0493	0,0602	0,021	0,182

5.2.3 Deployment Fredrik

Tabela 5.13 – Número de espalhadores da região da floresta Fredrik

Imagem	m_1	m_2	m_3	m_4	N
1	0,147	0,028	0,0067	0,00188	10
2	0,148	0,03	0,0078	0,0025	3,684
3	0,15	0,296	0,0071	0,002	10
4	0,138	0,026	0,006	0,0019	3,23
5	0,135	0,024	0,0053	0,0014	6,77
6	0,137	0,025	0,056	0,0015	7,478

Tabela 5.14 – Distância KS para a região da floresta do deployment Fredrik

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,148	0,0288	0,0624	0,0118	0,185
2	0,22076	0,0582	0,0577	0,0264	0,165
3	0,152	0,0259	0,066	0,0105	0,182
4	0,251	0,0666	0,0534	0,026	0,173
5	0,197	0,0373	0,062	0,017	0,191
6	0,182	0,0382	0,0606	0,0156	0,190

5.2.4 Deployment Adolf-Fredrik

Tabela 5.15 – Número de espalhadores da região da floresta

Imagem	m_1	m_2	m_3	m_4	N
1	0,153	0,03009	0,0073	0,00208	10,06
2	0,134	0,0254	0,0062	0,0019	3,075
3	0,16	0,033	0,0085	0,0025	1,563
4	0,136	0,0259	0,0064	0,002	3,018
5	0,136	0,0259	0,0064	0,002	3,018
6	0,134	0,024	0,00526	0,00138	7,37

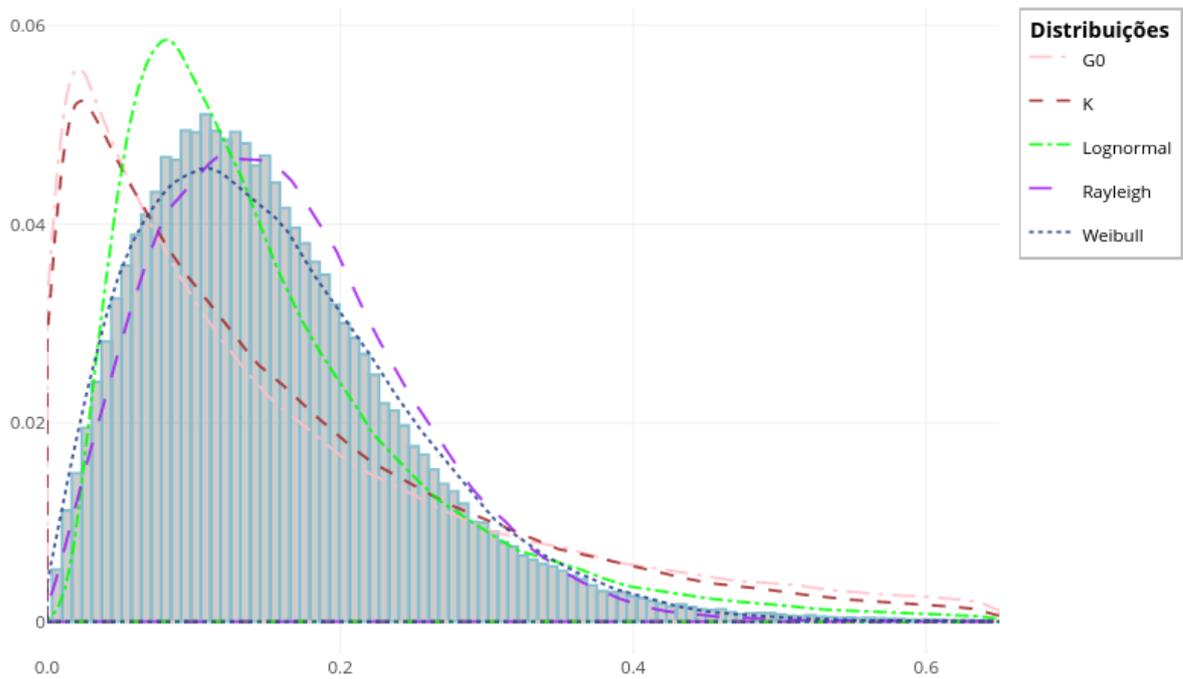
Tabela 5.16 – Distância KS para a região da floresta do deployment Adolf-Fredrik.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,156	0,0287	0,0633	0,0128	0,188
2	0,257	0,0682	0,057	0,0269	0,176
3	0,168	0,0236	0,069	0,0126	0,183
4	0,257	0,071	0,0538	0,0261	0,178
5	0,255	0,0712	0,0538	0,0261	0,18
6	0,191	0,0336	0,0621	0,017	0,183

5.2.5 Visualização das distribuições da região da floresta

O histograma dos dados da primeira imagem do *deployment* Karl é mostrado na Figura 5.5 e as curvas das PDFs ajustadas aos dados são exibidas com as cores e tipos de linhas indicadas na legenda do gráfico. A função de distribuição acumulada da amostra da região da floresta e das distribuições é mostrada na Figura 5.6.

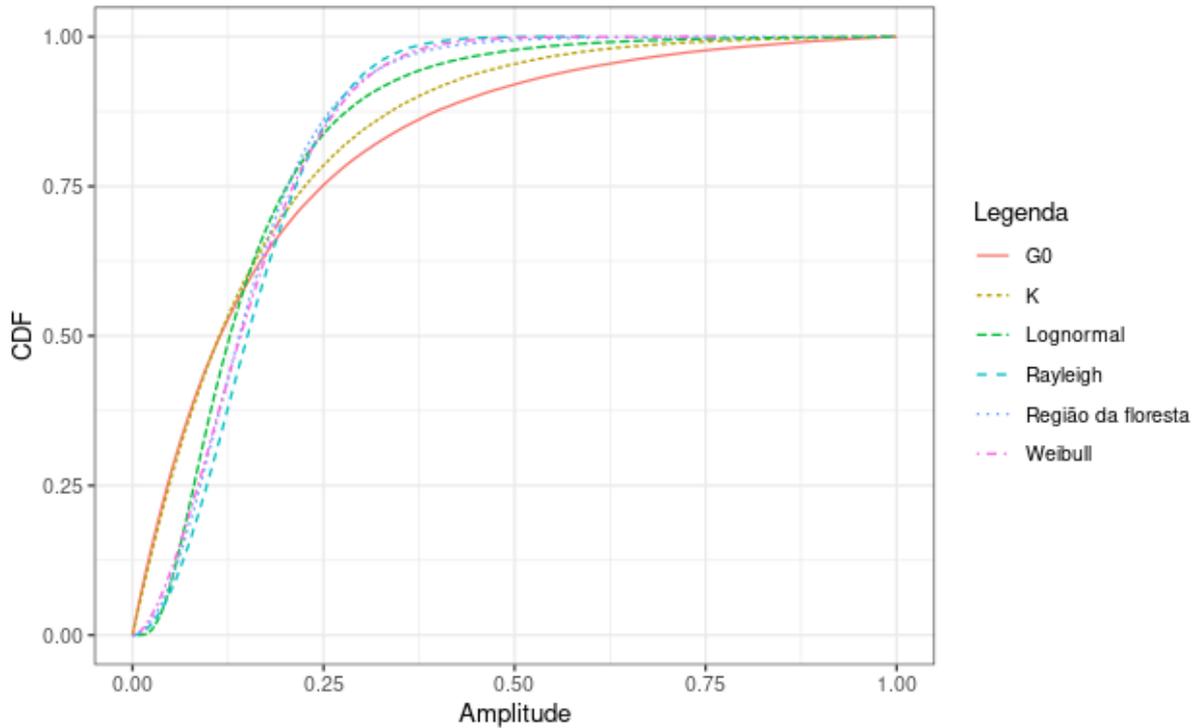
Figura 5.5 – Comparação das distribuições da região da floresta



Fonte: Próprio autor

O gráfico da distribuição empírica acumulada da Figura 5.6 da distribuição mostra que as distribuições mais similares a da região da floresta foram as distribuições Weibull e Rayleigh.

Figura 5.6 – CDF da região da floresta da imagem do deployment *sigismund*.



Fonte: Próprio autor

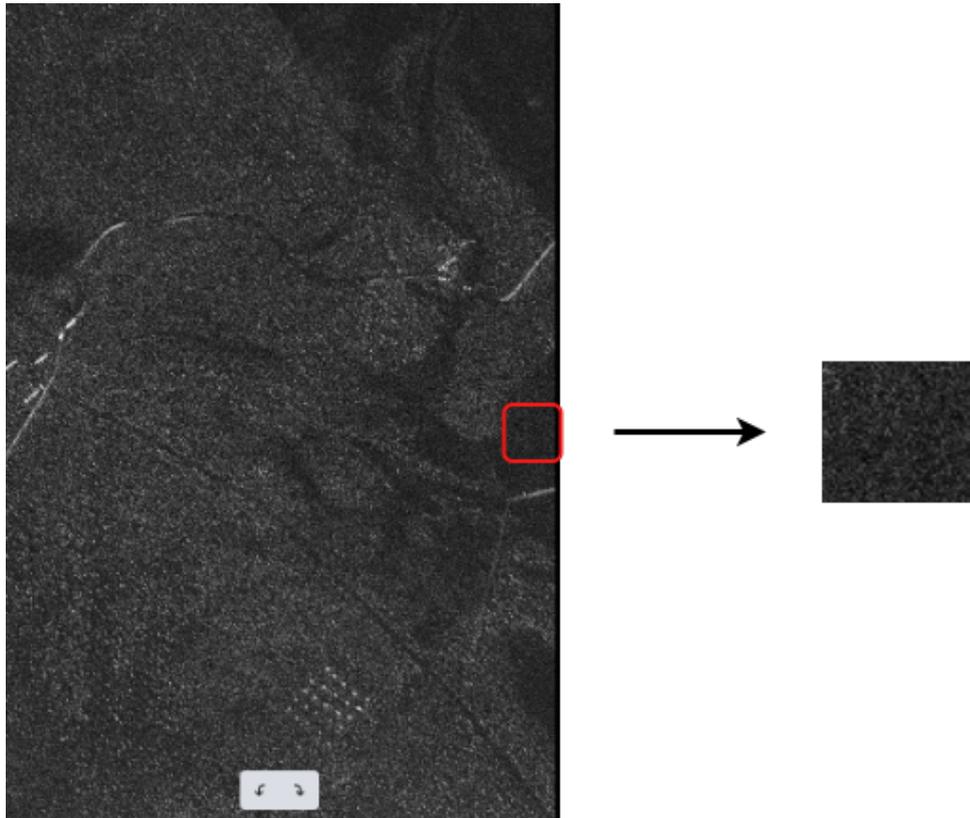
É interessante notar que ocorreu uma sobreposição no gráfico da ECDF da distribuição Weibull e dos dados da amostra. O teste *KS* basicamente faz a maior diferença absoluta da ECDF, esta sobreposição indica uma menor distância dessa métrica que é corroborado pelo resultado na Tabela 5.12, cujo resultado para a distância foi de 0,022, próximo do valor ótimo.

5.3 MODELAGEM ESTATÍSTICA DA REGIÃO DO LAGO

A região do lago foi selecionada após ter sido feito um recorte da região com base no mapa da região mostrado na Figura 4.1 do Capítulo 4. Essa etapa precisou ser feita pelo conjunto de dados não apresentar uma lista de referência com as coordenadas dessa região.

A Figura 5.7 apresenta a quinta imagem do *deployment* Adolf-Fredrik destacando a área analisada para cada imagem dos quatro *deployments*.

Figura 5.7 – Imagem do *deployment* Adolf-Fredrik para análise da região do lago.



Fonte: Próprio autor

5.3.1 Deployment Sigismund

Tabela 5.17 – Número de espalhadores da região do lago *deployment* Sigismund

Imagem	m_1	m_2	m_3	m_4	N
1	0,115	0,0175	0,0033	0,00075	6,44
2	0,115	0,0175	0,0033	0,00075	6,44
3	0,125	0,0205	0,0041	0,00098	9,22
4	0,109	0,018	0,006	0,0053	0,2163
5	0,113	0,0169	0,00316	0,00072	5,599
6	0,0988	0,0133	0,0023	0,0052	3,137

Tabela 5.18 – Distância KS das distribuições para a região do lago do *deployment* Sigismund

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,181	0,0298	0,0631	0,0161	0,188
2	0,179	0,0298	0,0632	0,0161	0,191
3	0,221	0,0224	0,0668	0,0116	0,191
4	0,415	0,094	0,062	0,051	0,187
5	0,170	0,033	0,064	0,0203	0,189
6	0,191	0,0492	0,058	0,0249	0,186

5.3.2 Deployment Karl

Tabela 5.19 – Número de espalhadores da região do lago do *deployment* Karl

Imagem	m_1	m_2	m_3	m_4	N
1	0,115	0,018	0,0033	0,0008	5,868
2	0,118	0,0208	0,00722	0,00586	0,2618
3	0,110	0,0160	0,00292	0,000656	5,503
4	0,110	0,0185	0,00701	0,00648	0,179
5	0,117	0,0181	0,00354	0,00084	5,412
6	0,104	0,0144	0,00256	0,000575	3,983

Tabela 5.20 – Distância KS das distribuições para a região do lago do *deployment* Karl.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,173	0,032	0,064	0,017	0,186
2	0,385	0,084	0,0642	0,0479	0,186
3	0,165	0,0295	0,0648	0,0164	0,192
4	0,428	0,1035	0,06	0,052	0,178
5	0,168	0,0378	0,0636	0,0192	0,187
6	0,159	0,0428	0,0626	0,0236	0,187

5.3.3 Deployment Fredrik

Tabela 5.21 – Número de espalhadores da região do lago do *deployment* Fredrik

Imagem	m_1	m_2	m_3	m_4	N
1	0,141	0,0276	0,0072	0,00277	1,845
2	0,140	0,0296	0,0098	0,00615	0,601
3	0,146	0,0294	0,0080	0,00315	1,841
4	0,136	0,0277	0,00897	0,00554	0,581
5	0,140	0,0296	0,00985	0,00615	0,601
6	0,144	0,0292	0,00814	0,00341	1,511

Tabela 5.22 – Distância KS das distribuições para a região do lago do *deployment* Fredrik.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0.1697	0,06373	0,0563	0,0253	0,178
2	0,2743	0,115	0,00485	0,00412	0,179
3	0,165	0,063	0,0565	0,00253	0,165
4	0,282	0,115	0,00483	0,004144	0,175
5	0,187	0,0073	0,00546	0,00271	0,175
6	0,182	0,00271	0,00545	0,00718	0,1753

5.3.4 Deployment Adolf-Fredrik

Tabela 5.23 – Número de espalhadores da região do lago do *deployment* Adolf-Fredrik

Imagem	m_1	m_2	m_3	m_4	N
1	0,118	0,0183	0,00351	0,00081	7,26
2	0,12	0,0202	0,0098	0,006153	0,601
3	0,124	0,0202	0,00403	0,000957	8,929
4	0,1126	0,0191	0,00689	0,00595	0,212
5	0,117	0,0191	0,00689	0,00595	0,2122
6	0,1065	0,0152	0,00273	0,0006137	4,534

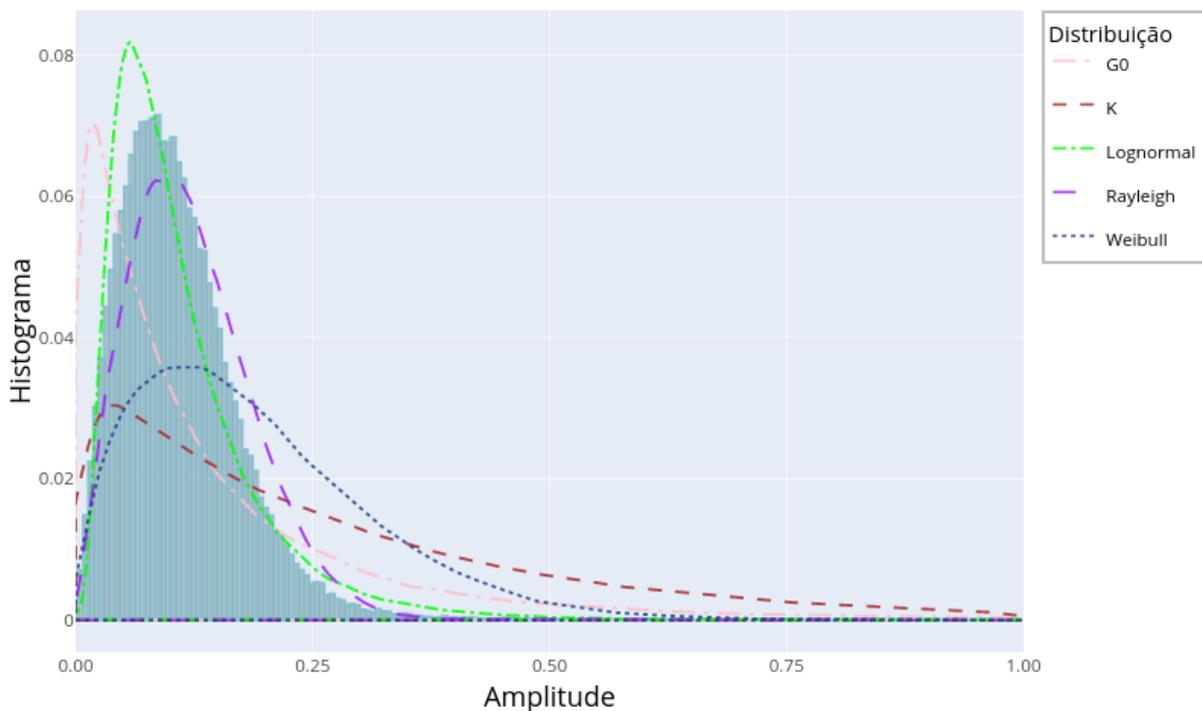
Tabela 5.24 – Distância KS das distribuições para a região do lago do *deployment* Adolf-Fredrik.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,197	0,0257	0,0671	0,0147	0,192
2	0,419	0,0973	0,0599	0,0516	0,182
3	0,2214	0,0231	0,0659	0,0125	0,193
4	0,412	0,0923	0,0611	0,0525	0,186
5	0,411	0,0923	0,0611	0,0525	0,187
6	0,149	0,03845	0,0632	0,01966	0,188

5.3.5 Visualização das distribuições da região do lago

A Figura 5.8 mostra o histograma em azul e as curvas das distribuições para a primeira imagem do *deployment* Adolf-Fredrik e a função distribuição acumulada de cada distribuição e dos dados são exibidos na Figura 5.9.

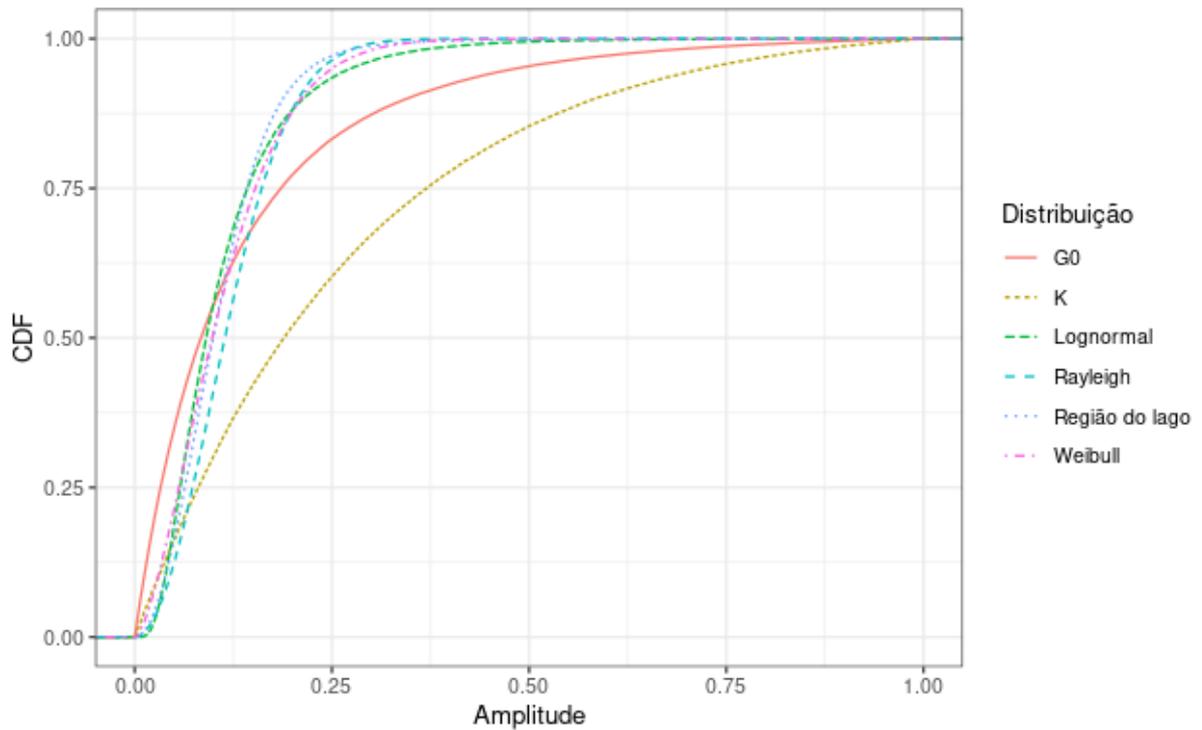
Figura 5.8 – Comparação das distribuições da região do lago



Fonte: Próprio autor

O gráfico da distribuição empírica acumulada da Figura 5.9 da distribuição mostra de forma clara que as distribuições que melhor descreveram a amostra da região da floresta foram as distribuições Weibull e Rayleigh.

Figura 5.9 – CDF da região do lago da primeira imagem do deployment *Adolf-Fredrik*.

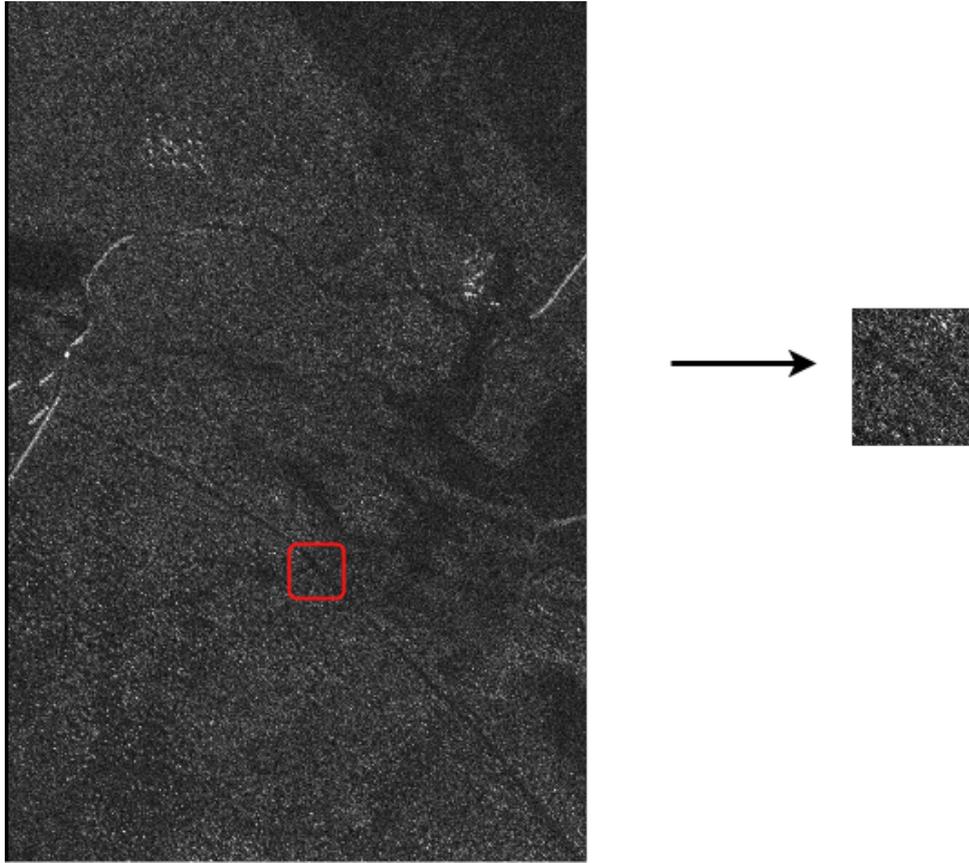


Fonte: Próprio autor

5.4 MODELAGEM ESTATÍSTICA DA REGIÃO DA ESTRADA

A região da estrada foi selecionada após uma inspeção com base no mapa da região mostrado na Figura 4.1 do Capítulo 4. A Figura 5.10 mostra uma das imagens dessa região que foi usada neste estudo. É importante destacar que esta região é envolta pela floresta, o que torna mais difícil a seleção da área de interesse.

Figura 5.10 – Imagem do *deployment* Karl para análise da região da estrada.



Fonte: Próprio autor

5.4.1 Deployment Sigismund

Tabela 5.25 – Número de espalhadores da região da estrada para o *deployment* Sigismund.

Imagem	m_1	m_2	m_3	m_4	N
1	0,154	0,0323	0,00867	0,00287	4,0391
2	0,154	0,0323	0,00867	0,002867	4,039
3	0,159	0,0344	0,00951	0,00324	4,145
4	0,1543	0,0352	0,0110	0,00442	1,927
5	0,160	0,0345	0,00936	0,00309	5,0757
6	0,145	0,029	0,00736	0,002304	4,056

Tabela 5.26 – Distância KS das distribuições para a região da estrada do *deployment* Sigismund

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,175	0,058	0,0572	0,0239	0,179
2	0,175	0,0580	0,0572	0,0239	0,182
3	0,178	0,0570	0,0570	0,0218	0,178
4	0,152	0,118	0,0436	0,0426	0,161
5	0,2014	0,0489	0,0583	0,0195	0,181
6	0,167	0,0602	0,059	0,0235	0,179

5.4.2 Deployment Karl

Tabela 5.27 – Número de espalhadores da região da estrada para o *deployment* Karl.

Imagem	m_1	m_2	m_3	m_4	N
1	0,156	0,0336	0,00925	0,00315	3,806
2	0,156	0,0336	0,00925	0,00315	3,806
3	0,152	0,0316	0,0084	0,00280	3,718
4	0,157	0,036	0,0116	0,0047	1,987
5	0,172	0,0386	0,01041	0,00326	14,97
6	0,152	0,0314	0,00829	0,00265	4,126

Tabela 5.28 – Distância KS das distribuições para a região da estrada do *deployment* Karl.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,170	0,0643	0,05391	0,0247	0,176
2	0,170	0,0642	0,0539	0,0247	0,175
3	0,168	0,0627	0,544	0,0253	0,18
4	0,141	0,122	0,0423	0,04289	0,162
5	0,312	0,0241	0,0639	0,00672	0,188
6	0,175	0,0565	0,057	0,0223	0,18

5.4.3 Deployment Fredrik

Tabela 5.29 – Número de espalhadores da região da estrada para o *deployment* Fredrik.

Imagem	m_1	m_2	m_3	m_4	N
1	0,1497	0,0306	0,00799	0,00256	4,133
2	0,158	0,0371	0,0119	0,00485	1,978
3	0,153	0,032	0,00856	0,00284	3,853
4	0,154	0,03501	0,01098	0,004397	1,924
5	0,1579	0,0337	0,00911	0,003027	4,544
6	0,161	0,0351	0,00969	0,00329	4,501

Tabela 5.30 – Distância KS das distribuições para a região da estrada do *deployment* Fredrik.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,172	0,0591	0,0567	0,023	0,178
2	0,142	0,122	0,0432	0,0414	0,158
3	0,168	0,0578	0,0057	0,023	0,181
4	0,149	0,124	0,0429	0,0424	0,158
5	0,19	0,0506	0,0585	0,0207	0,182
6	0.19147	0,05162	0,0587	0,0219	0,182

5.4.4 Deployment Adolf-Fredrik

Tabela 5.31 – Número de espalhadores da região do lago do *deployment* Adolf-Fredrik.

Imagem	m_1	m_2	m_3	m_4	N
1	0,152	0,0317	0,00839	0,00274	4,1134
2	0,1526	0,0346	0,01076	0,00428	1,915
3	0,1641	0,03663	0,01034	0,00360	4,404
4	0,1565	0,0362	0,01142	0,00462	1,974
5	0,1565	0,0362	0,01143	0,00462	1,974
6	0,1522	0,0317	0,00839	0,002741	4,114

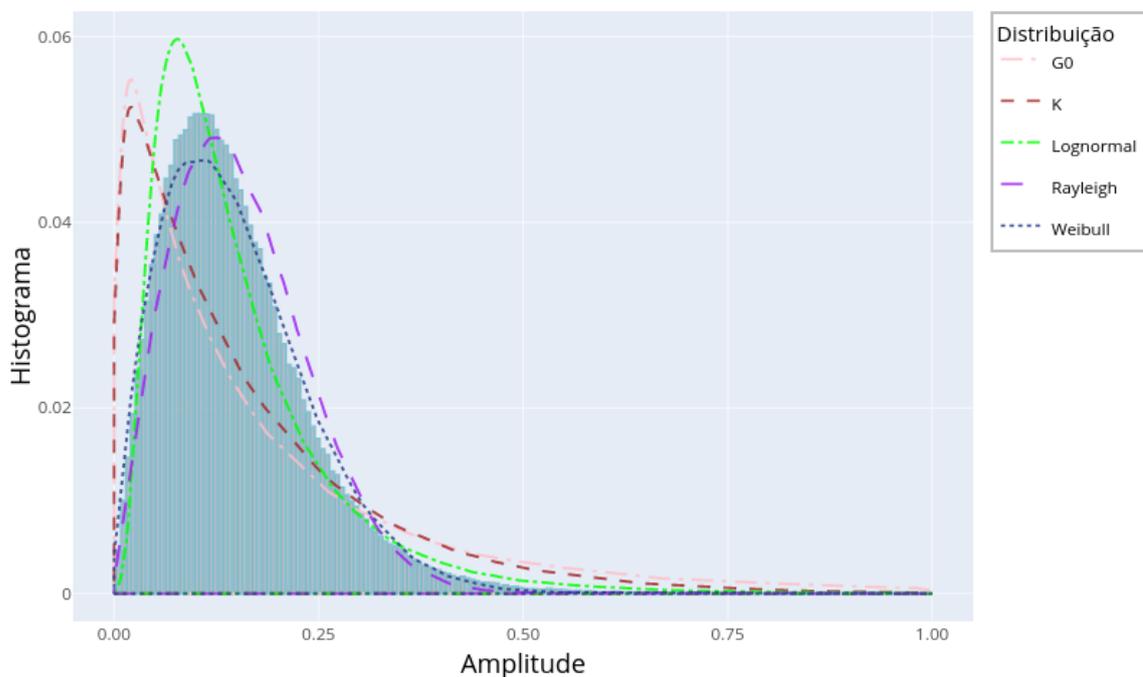
Tabela 5.32 – Distância KS das distribuições para a região da estrada do deployment Adolf-Fredrik.

Imagem	K	Rayleigh	LogNormal	Weibull	G0
1	0,172	0,0582	0,0561	0,023	0,18
2	0,138	0,122	0,0435	0,0421	0,158
3	0,192	0,0506	0,0576	0,023	0,18
4	0,143	0,119	0,0428	0,0413	0,16
5	0,142	0,119	0,0428	0,0413	0,159
6	0,174	0,0582	0,0561	0,023	0,178

5.4.5 Visualização das distribuições da região da estrada

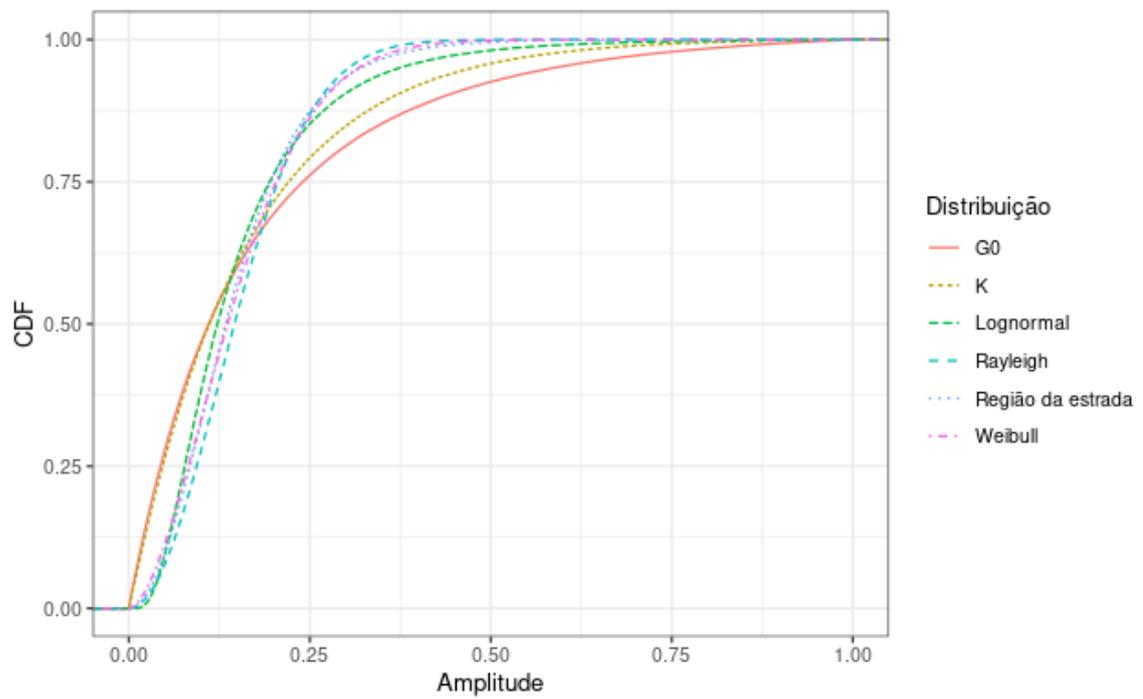
A Figura 5.11 mostra as cinco distribuições da primeira imagem do *deployment* Fredrik para a região da estrada e o seu histograma em azul. As funções de distribuição acumulada dos dados e das distribuições ajustadas são mostradas na Figura 5.12.

Figura 5.11 – Comparação das distribuições da região da estrada.



Fonte: Próprio autor

As distribuições que melhor representaram os dados da estrada foram as distribuições Weibull e Rayleigh, conforme pode ser visualizado na função de distribuição acumulada da Figura 5.12 essas regiões tiveram o resultado mais próximo ao da amostra.

Figura 5.12 – CDF da região da estrada da primeira imagem do deployment *Fredrik*.

Fonte: Próprio autor

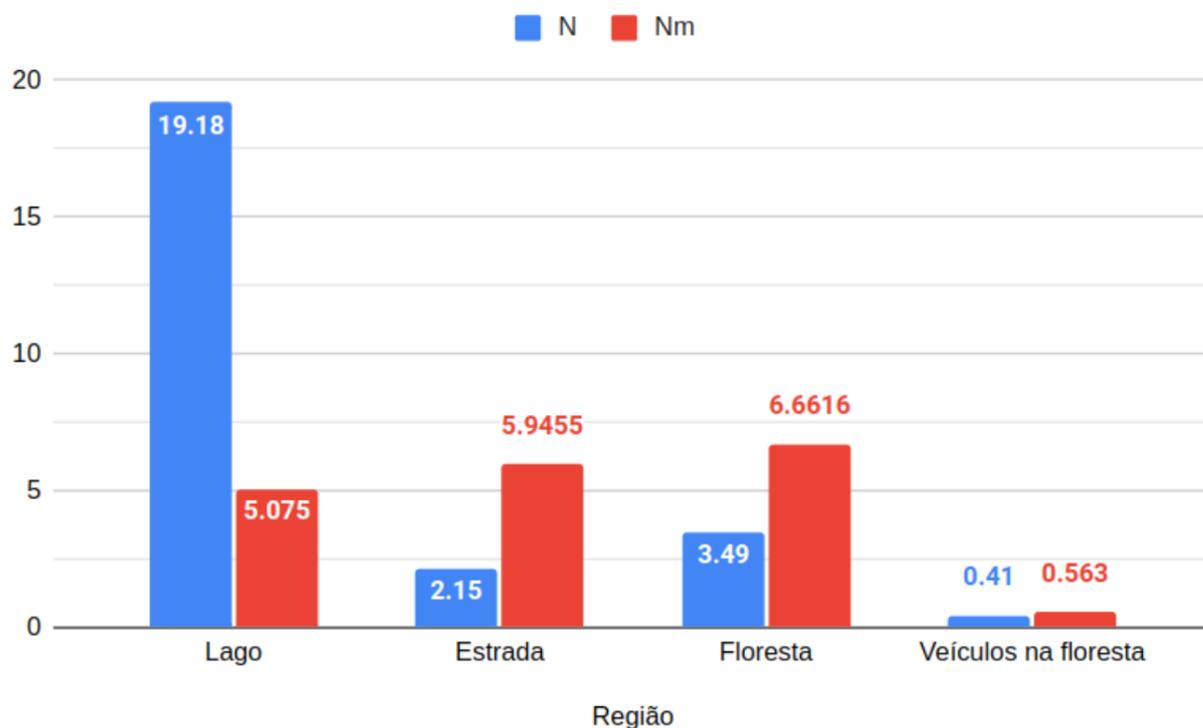
6 ANÁLISE DOS RESULTADOS

Este capítulo apresenta uma análise dos número de espalhadores e do ajuste das distribuições para cada região. Destaca-se que para facilitar a comparação com o estudo (XU, 2022) optou-se por realizar a média. Dessa forma, são analisados os número de espalhadores médios de cada região e a média da distância KS .

6.1 NÚMERO DE ESPALHADORES

A Figura 6.1 apresenta a comparação feita com N sendo o número de espalhadores do estudo original e Nm a média dos espalhadores para cada região encontrados neste trabalho de conclusão de curso.

Figura 6.1 – Comparação do número de espalhadores N do estudo de referência e Nm sendo o número de espalhadores médio estimados no estudo.



Fonte: Próprio autor

Nota-se que os valores mais próximos foram calculados na região dos veículos na floresta, esta região é caracterizada por poucos espalhadores com maior amplitude, corroborado por ambos resultados mostrados na Figura 6.1, com $N = 0,41$ e $Nm = 0,563$.

A maior diferença foi encontrada no número de espalhadores da região do lago com $N = 19,18$ da região do lago, bastante superior a média de $Nm = 5,075$ estimada

neste trabalho. A região do lago é caracterizada por apresentar uma constante dielétrica baixa, no qual as ondas incidentes de energia penetram a camada da superfície e ocorre o espalhamento volumétrico. Em geral essa região apresenta um número de espalhadores elevado devido a essa característica, porém o número de espalhadores estimado foi inferior as encontradas nas regiões da floresta e estrada.

O espalhamento na região com vegetação é influenciado por fatores como a orientação dos galhos, teor de umidade, presença de folhas e suas proporções nas camadas do dossel. Nessa região ocorreu uma maior variação entre os valores calculados devidos as características distintas entre as regiões capturadas pelo sistema SAR. A região da estrada apresentou um valor próximo de espalhadores a essa região, devido a presença da floresta na área de estudo. O número de espalhadores estimados para a região da floresta e estrada foram maiores do que os apresentados no estudo de (XU, 2022).

De acordo com (XU, 2022), o modelo do speckle completamente desenvolvido é atingido desde que $N > 8$. Embora no estudo de referência essa condição tenha sido atendida para a região do lago, nenhuma das regiões estudadas satisfaz essa condição nas simulações realizadas.

6.2 MODELAGEM DAS DISTRIBUIÇÕES

A Tabela 6.1 mostra o resultado ajuste das distribuições obtido fazendo a média da distância KS para cada região analisada, enquanto que a Tabela 6.2 mostra os valores do estudo de referência apresentados para comparação.

Tabela 6.1 – Média da distância KS das distribuições do trabalho de conclusão de curso.

Região	K	Lognormal	Rayleigh	Weibull	G0
Veículos na floresta	0,336	0,049	0,1601	0,0547	0,184
Floresta	0.247	0,0567	0,0904	0,0626	0,1783
Lago	0,246	0,0528	0,0537	0,0236	0,184
Estrada	0,173	0,0915	0,105	0,0278	0,174

Tabela 6.2 – Ajuste de distribuições usando a distância KS do estudo de referência.

Região	K	Lognormal	Rayleigh	Weibull
Veículos na floresta	0,279	0,063	0,226	0,081
Floresta	0,144	0,078	0,054	0,041
Lago	0,144	0,103	0,023	0,036
Estrada	0,037	0,040	0,176	0,045

Conforme apresentado na Tabela 6.1 as distribuições Weibull e Lognormal conseguiram descrever as regiões com diferentes características de espalhamento conforme apresentado pela distância KS . Ambas distribuições pertencem a categoria de modelos empíricos de distribuição única descrita no Capítulo 3, que modelam a distribuição sem necessariamente considerar a interpretação física na formação da imagem (YUE et al., 2021).

No estudo de referência para região do lago e estrada as distribuições Rayleigh e K apresentaram, respectivamente, os melhores valores da distância KS conforme mostra a Tabela 6.2. A Tabela 6.1 mostra que esse resultado não foi atingido nas simulações feitas, onde a distribuição Weibull apresentou o melhor ajuste aos dados.

A distribuição $G0$ é interessante para modelar regiões heterogêneas e de acordo com (FRERY; WU; GOMEZ, 2022) consegue aproximar a distribuição K . Na Tabela 6.1 mostra que em quase todas as regiões essa distribuição apresentou um resultado superior levando em conta a distância KS em relação a distribuição K .

A distribuição Rayleigh é usada principalmente em regiões homogêneas como é o caso da região do lago. O valor da distância KS foi próximo ao obtido no estudo, o qual esta distribuição obteve a menor distância observada pela métrica KS . É interessante observar que em outras regiões, principalmente a região dos veículos na floresta esta distribuição apresentou um ajuste inferior. Atrelado a isso está o fato de que a região apresenta os alvos que representam sinais de retorno de maior amplitude, no qual o modelo de Rayleigh não é adequado para modelar. Destaca-se que em nenhuma região a distribuição K obteve o melhor ajuste conforme pode ser observado na Tabela 6.1.

7 CONCLUSÃO

Este estudo com base no artigo no artigo (XU, 2022) realizou uma abordagem quantitativa do cálculo do número de espalhadores de 24 imagens SAR VHF do conjunto de dados CARABAS-II . A partir dos resultados apresentados neste trabalho de conclusão de curso verifica-se que para a aplicação de detecção de alvos a distribuição que melhor se ajusta a região dos veículos na floresta foi a distribuição Lognormal. A princípio essa distribuição irá produzir uma maior assertividade na detecção de número de alvos, uma vez que melhor representa o sinal de retorno.

Os estudo comparativo das distribuições mostrou que os modelos empíricos para imagens SAR VFH conseguem descrever de maneira satisfatória regiões com diferentes tipos de espalhamento, sendo úteis para a modelagem de imagens SAR VHF. Destaca-se que o cálculo do número de espalhadores foi inconclusivo para as demais regiões, com exceção dos veículos na floresta. Apesar de constituir uma ferramenta interessante para o estudo de imagens SAR, este método pode ainda ser refinado para adequar as variações do sinal de retorno das regiões da floresta, lago e estrada. Estudos posteriores podem ser desenvolvidos nesse aspecto, inicialmente testando o método com uma nova base de dados e avaliando a variação de N para cada região.

REFERÊNCIAS BIBLIOGRÁFICAS

AGER, T. An Introduction to Synthetic Aperture Radar Imaging. **Oceanography**, v. 26, n. 2, jun. 2013. ISSN 10428275. Disponível em: <<https://tos.org/oceanography/article/an-introduction-to-synthetic-aperture-radar-imaging>>.

BATISTA, J. L. F. Verossimilhança e Máxima Verossimilhança. p. 27, 2009.

CENTENO, J. A. S. **Sensoriamento Remoto**. 2019. Disponível em: <https://docs.ufpr.br/~centeno/p_sr1/pdf/aula07.pdf>.

DEVORE, J.; SILVA, J. da. **Probabilidade e estatística: para engenharia e ciências**. Pioneira Thomson Learning, 2006. ISBN 9788522104598. Disponível em: <<https://books.google.com.br/books?id=1JOIPgAACAAJ>>.

FRERY, A.; WU, J.; GOMEZ, L. **SAR Image Analysis - A Computational Statistics Approach: With R Code, Data, and Applications**. Wiley, 2022. ISBN 9781119795292. Disponível em: <<https://books.google.com.br/books?id=jNguzgEACAAJ>>.

GAO, G. Statistical Modeling of SAR Images: A Survey. **Sensors (Basel, Switzerland)**, v. 10, n. 1, p. 775–795, jan. 2010. ISSN 1424-8220. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3270869/>>.

INPE. **Teoria : Sistemas de Radar**. 2019. Disponível em: <<http://www.dpi.inpe.br/spring/teoria/radar/radar.htm>>.

ISKANDER, D.; ZOUBIR, A. Estimation of the parameters of the K-distribution using higher order and fractional moments [radar clutter]. **IEEE Transactions on Aerospace and Electronic Systems**, v. 35, n. 4, p. 1453–1457, out. 1999. ISSN 1557-9603. Conference Name: IEEE Transactions on Aerospace and Electronic Systems.

KOO, Y. K. C. a. V. An Introduction to Synthetic Aperture Radar (SAR). **Progress In Electromagnetics Research B**, v. 2, p. 27–60, 2008. ISSN 1937-6472. Publisher: EMW Publishing. Disponível em: <<http://www.jpier.org/PIERB/pier.php?paper=07110101>>.

LEE, J.-S.; POTTIER, E. (Ed.). **Polarimetric Radar Imaging: From Basics to Applications**. Boca Raton: CRC Press, 2017. ISBN 978-1-315-21933-2.

LU, J. Introduction. In: _____. **Design Technology of Synthetic Aperture Radar**. [S.l.: s.n.], 2019. p. 1–13.

LUNDBERG, M. et al. A challenge problem for detection of targets in foliage. In: ZELNIO, E. G.; GARBER, F. D. (Ed.). Orlando (Kissimmee), FL: [s.n.], 2006. p. 62370K. Disponível em: <<http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.663594>>.

MARWAN, Y. Tutorial, **Synthetic Aperture Radar (SAR): Principles and Applications**. University of Agronomic Science and Veterinary Medicine in Bucharest: [s.n.], 2015. Disponível em: <https://eo4society.esa.int/wp-content/uploads/2021/02/D1T2a_LTC2015_Younis.pdf>.

MAÎTRE, H. **Processing of Synthetic Aperture Radar (SAR) Images | Wiley**. 2013. Disponível em: <<https://www.wiley.com/en-sg/Processing+of+Synthetic+Aperture+Radar+%28SAR%29+Images-p-9781848210240>>.

MEDEIROS, F. N. S. d. **Filtragem adaptativa de imagens de radar de abertura sintética utilizando a abordagem maximum a posteriori**. dez. 1999. Tese (text) — Universidade de São Paulo, dez. 1999. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/76/76132/tde-27112014-164749/>>.

MOREIRA, A. et al. A tutorial on synthetic aperture radar. **IEEE Geoscience and Remote Sensing Magazine**, v. 1, n. 1, p. 6–43, mar. 2013. ISSN 2168-6831. Conference Name: IEEE Geoscience and Remote Sensing Magazine.

NICOLAS, J.-M.; ANFINSEN, S. Introduction to second kind statistics: Application of log-moments and log-cumulants to analysis of radar images. **Traitement du Signal**, v. 3, 01 2002.

OLIVER, C.; QUEGAN, S. Understanding synthetic aperture radar images. 01 2004.

OTE. **Radar de Abertura Sintética (SAR - Synthetic Aperture Radar) para satélites de pequeno porte**. 2019. Disponível em: <https://www.cgee.org.br/documents/10195/734063/5478.pdf_Documentos+Estrat%C3%A9gicos+para+o+Setor+Espacial+-+Radar+de+Abertura+Sint%C3%A9tica+-+SAR+-+Synthetic+Aperture+Radar+-+para+sat%C3%A9lites+de+pequeno+porte_vf.pdf/dbf7363b-75fd-4137-afd7-98efa2f780d1?version=4.0>.

RICCI, V. Fitting distributions with r. 01 2005.

TREVETT, J. W. other, **Imaging radar for resources surveys**. 2008. Archive Location: world Publisher: Taylor & Francis Group. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/10106048709354112>>.

UFS. **MEDIDAS DE DISPERSÃO, ASSIMETRIA E CURTOSE**. 2019. Disponível em: <https://cesad.ufs.br/ORBI/public/uploadCatalogo/09215427022012Bioestatistica_Aula_03.pdf>.

ULABY, F. T. et al. **Microwave Radar and Radiometric Remote Sensing**. [S.l.]: University of Michigan Press, 2014. Google-Books-ID: y6pZngEACAAJ. ISBN 978-0-472-11935-6.

VIEIRA, S. **Bioestatística: Tópicos avançados - testes não paramétricos, testes diagnósticos, medidas de associação e concordância**. Elsevier Editora Ltda., 2018. ISBN 9788535289824. Disponível em: <<https://books.google.com.br/books?id=UodjDwAAQBAJ>>.

XU, Z. Wavelength-resolution sar speckle model. **IEEE Geoscience and Remote Sensing Letters**, v. 19, p. 1–5, 2022.

YUE, D.-X. et al. Synthetic Aperture Radar Image Statistical Modeling: Part One-Single-Pixel Statistical Models. **IEEE Geoscience and Remote Sensing Magazine**, v. 9, n. 1, p. 82–114, mar. 2021. ISSN 2168-6831. Conference Name: IEEE Geoscience and Remote Sensing Magazine.